

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2004, 07:40:11 ; Search time 2570.42 Seconds
(without alignments)
9398.725 Million cell updates/sec

Title: US-09-864-675-1
Perfect score: 994
Sequence: 1 atgaggcgcgcacccggcccc.....caccttggatttgaattaaa 994

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*

28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8							
Result	Query						
No.	Score	Match	Length	DB	ID	Description	
	1	674	67.8	805	12	BI918620	BI918620 603176570
	2	586	59.0	1047	12	BM914622	BM914622 AGENCOURT
c	3	467.2	47.0	1041	12	BI412864	BI412864 602988202
	4	467	47.0	524	13	BX281777	BX281777 BX281777
	5	404	40.6	549	9	AA706226	AA706226 ah28a07.s
	6	395.6	39.8	412	9	AI041451	AI041451 ow36c02.s
	7	381	38.3	488	4	BX529505	Bx529505 RZPD Mus
	8	363.8	36.6	427	10	BF108794	BF108794 7152g03.x
c	9	317.4	31.9	949	12	BI410828	BI410828 602963734
	10	272.6	27.4	795	12	BI651936	BI651936 603298677
	11	234.4	23.6	333	10	BE983573	BE983573 UI-M-CG0p
	12	232.2	23.4	529	9	AW476657	AW476657 uq79e01.y
	13	224	22.5	297	9	AA772412	AA772412 ai44e12.s
c	14	180	18.1	362	13	BX089049	BX089049 BX089049
	15	171.8	17.3	256	9	AW762061	AW762061 ur53c01.y
	16	154	15.5	657	13	BQ078813	BQ078813 fy81c06.y
	17	151.8	15.3	259	10	BE648780	BE648780 UI-M-BH2.
	18	151.4	15.2	327	9	AA968077	AA968077 uh09h01.r
c	19	145	14.6	539	9	AL918370	AL918370 AL918370
	20	109	11.0	458	9	AI152190	AI152190 ud18h10.r
	21	105.2	10.6	750	29	BZ847665	BZ847665 CH240_239
	22	101	10.2	493	28	BH057870	BH057870 RPCI-24-9
c	23	100.4	10.1	481	28	AZ987593	AZ987593 2M0270P10
	24	90.2	9.1	243	10	BB570162	BB570162 BB570162
c	25	83.6	8.4	512	9	AI073386	AI073386 oo13d06.x
	26	81.2	8.2	167	9	AI836531	AI836531 UI-M-AP0-
	27	76.6	7.7	477	10	BE984041	BE984041 UI-M-CG0p
	28	68.4	6.9	769	12	BI413085	BI413085 602990205
	29	67.2	6.8	765	12	BI522417	BI522417 603175321
	30	67.2	6.8	818	12	BI651928	BI651928 603298668
	31	66	6.6	751	29	CNS04J6G	AL293137 Tetraodon
	32	64.8	6.5	321	10	BE983721	BE983721 UI-M-CG0p
	33	64.4	6.5	538	9	AL925790	AL925790 AL925790
c	34	61	6.1	491	9	AL909688	AL909688 AL909688
	35	60.4	6.1	413	14	N62228	N62228 yz63c08.sl
c	36	60.2	6.1	322	9	AL909689	AL909689 AL909689
	37	58.4	5.9	685	14	CA351220	CA351220 622234 NC
	38	53	5.3	1630	11	AK051824	AK051824 Mus muscu
	39	51.6	5.2	925	29	CNS0091P	AL053013 Drosophil
c	40	51.6	5.2	982	13	BX415111	BX415111 BX415111
	41	50.8	5.1	647	12	BI960178	BI960178 HVSMEn002
	42	50.4	5.1	251	9	AW045376	AW045376 UI-M-BH1-
	43	49.8	5.0	473	12	BI666105	BI666105 603287281
	44	49.8	5.0	658	14	CB059196	CB059196 NISC_jx12
	45	49.8	5.0	705	12	BI662853	BI662853 603286287

ALIGNMENTS

RESULT 1

BI918620

LOCUS BI918620 805 bp mRNA linear EST 16-OCT-2001

DEFINITION 603176570F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5240969 5', mRNA sequence.

ACCESSION BI918620

VERSION BI918620.1 GI:16182295

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 805)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM11607 row: k column: 18

High quality sequence start: 2

High quality sequence stop: 778.

FEATURES

source

Location/Qualifiers

1. .805

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5240969"

/lab_host="DH10B"

/clone_lib="NIH_MGC_121"

/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH_MGC Library."

BASE COUNT 169 a 243 c 263 g 130 t

ORIGIN

Query Match 67.8%; Score 674; DB 12; Length 805;

Best Local Similarity 98.7%; Pred. No. 3.6e-156;

Matches 732; Conservative 0; Mismatches 5; Indels 5; Gaps 5;

Qy 1 ATGAGGCGCGACCCGGCCCCCGGC-TTCTCCATGCTGCTCTTCGGTGTGTGCTCGCCTG 59
 |||
 Db 64 ATGAGGCGCGACCCGGCCCCCGGCGTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCCTG 123
 Qy 60 CTACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGA 119
 |||
 Db 124 CTACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGA 183
 Qy 120 GGGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCC 179
 |||
 Db 184 GGGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCC 243
 Qy 180 GCCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGG 239
 |||
 Db 244 GCCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGG 303
 Qy 240 GGGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCA 299
 |||
 Db 304 GGGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCA 363
 Qy 300 GCGCTACATCTTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTG- 358
 |||
 Db 364 GCGCTACATCTTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGT 423
 Qy 359 CCCCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTG 418
 |||
 Db 424 CCCCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTG 483
 Qy 419 ACTGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGA 478
 |||
 Db 484 ACTGCGCCACCCGGCCCCAAGTTGAAGAAGATGACGAGCCAGACGGGACAGGTGGGTGAGA 543
 Qy 479 AGCAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCA 538
 |||
 Db 544 AGCAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCA 603
 Qy 539 AGGATGGCAAGGAGCTCAACCG-CAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGA 597
 |||
 Db 604 AGGATGGCAAGGAGCTCAACCGTCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGA 663
 Qy 598 AAGAACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGA-GTATGTCTG 656
 |||
 Db 664 AAGAACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGAGGTATGTCTG 723
 Qy 657 CGAGG-CCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCG 715
 |||
 Db 724 CGAGGCCCAGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGGT 783
 Qy 716 TGAGCACCACCCTGTCATCCTG 737
 |||
 Db 784 TGAGCACCACCTGTCATCCTG 805

RESULT 2
 BM914622
 LOCUS

BM914622

1047 bp

mRNA

linear

EST 12-MAR-2002

DEFINITION AGENCOURT_6615334 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:5480308
5', mRNA sequence.

ACCESSION BM914622

VERSION BM914622.1 GI:19365001

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1047)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. Mark Watson

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLCM2002 row: p column: 05

High quality sequence stop: 541.

FEATURES

source

Location/Qualifiers

1. .1047

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5480308"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_113"

/note="Organ: spleen; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

BASE COUNT 263 a 347 c 254 g 183 t

ORIGIN

Query Match 59.0%; Score 586; DB 12; Length 1047;

Best Local Similarity 98.2%; Pred. No. 2.4e-134;

Matches 603; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

Qy 272 GCTCCTGTGTGCCGCTCGAAAGGAACCAGCGCTACATCTTTTCTGGAGCCCACGGAAC 331
|||||

Db 1 GCTCCTGTGTGCCGCTCGAAAGGAACCAGCGCTACATCTTTTCTGGAGCCCACGGAAC 60

Qy 332 AGCCCTTAGTCTTTAAGACGGCCTTTGCCCCCTCGATACCAACGGCAAAAATCTCAAGA 391
|||||

Db 61 AGCCCTTAGTCTTTAAGACGGCCTTTGCCCCCTCGATACCAACGGCAAAAATCTCAAGA 120

Qy 392 AAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCCGGCCCAAGTTGAAGAAGATGA 451
|||||

Db 121 AAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCCGGCCCAAGTTGAAGAAGATGA 180

Qy 452 AGAGCCAGACGGGACAGGTGGGTGAGAAGCAATCGCTGAAGTGTGAGGCAGCAGCCGGTA 511
 |||

Db 181 AGAGCCAGACGGGACAGGTGGGTGAGAAGCAATCGCTGAAGTGTGAGGCAGCAGCCGGTA 240

Qy 512 ATCCCCAGCCTTCCTACCGTTGGTTCAAGGATGGCAAGGAGCTCAACCGCAGCCGAGACA 571
 |||

Db 241 ATCCCCAGCCTTCCTACCGTTGGTTCAAGGATGGCAAGGAGCTCAACCGCAGCCGAGACA 300

Qy 572 TTCGCATCAAATATGGCAACGGCAGAAAGAACTCACGACTACAGTTCAACAAGGTGAAGG 631
 |||

Db 301 TTCGCATCAAATATGGCAACGGCAGAAAGAACTCACGACTACAGTTCAACAAGGTGAAGG 360

Qy 632 TGGAGGACGCTGGGGAGTATGTCTGCGAGGCCGAGAACATCCTGGGGAAGGACACCGTCC 691
 |||

Db 361 TGGAGGACGCTGGGGAGTATGTCTGCGAGGCCGAGAACATCCTGGGGAAGGACACCGTCC 420

Qy 692 GGGGCCGGCTTTACGTCAACAGCGTGAGCACCACCCTGTCATCCTGGTCGGGGCAGCCCC 751
 |||

Db 421 GGGGCCGGCTTTACGTCAACAGCGTGAGCACCACCCTGTCATCCTGGTCGGGGCAGCCCC 480

Qy 752 GGAAGTGCAACGAGACAGCCAAGTCCTATTGCGTCAATGGAGGCGTCTGCTACTACATCG 811
 |||

Db 481 GGAAGTGCAACGAGACAGCCAAGTCCTATTGCGTCAATGGAGGCGCTGCTACTACATCG 540

Qy 812 AGGGCATCAACCAGCTCTCCTGCAAATGTCCAAATGGATTCTTCGGACAGAGATGTTT-G 870
 |||

Db 541 AGGCCATCAATCAGCTTTCCTGCAAATGTCCCAATGGATTCTTCGGACCAACATGTTTGG 600

Qy 871 GAGAAACTGCCTTT 884
 |||

Db 601 GAGAAACTGCCCTT 614

RESULT 3

BI412864/c

LOCUS BI412864 1041 bp mRNA linear EST 14-AUG-2001

DEFINITION 602988202F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5144016 5', mRNA sequence.

ACCESSION BI412864

VERSION BI412864.1 GI:15173787

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1041)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

High quality sequence stop: 645.

ORIGIN

Matches 563; Conservative 0; Mismatches 78; Indels 9; Gaps 4;

Qy	165	CAGCACCCGAGAGCCGCCCGCCTCGGGTCGGGT-----GGCGTTGGTAAGGTGCTGGACA	220
Db	656	CACCTCGAGATGCCCGCCCGCCTCGGGTTCGGTTGGCGTCTTGGTGAAAGGTGCTGGACA	597
Qy	221	AGTGGCCG--CTCCGGAGCGGGGGGCTGCAGCGCAGCAGGTGATCAGCGTGGGCTCCTG	278
Db	596	AGTTGCCGGCTCCCGGATCGGGGGGCTGCAGCGCAGCAGGTGATCAGCGTGGGCTCCTG	537
Qy	279	TGTGCCGCTCGAAAGGAACCAGCGCTACATCTT-TTTCCTGGAGCCACGGAACAGCCCT	337
Db	536	TGCGCCGCTCGAAAGGAACCAGCGCTACATCTTGTTTCCTGGAGCCACCGAGCAGCCCT	477
Qy	338	TAGTCTTTAAGACGGCCTTTGCCCCCTCGATACCAACGGCAAAATCTCAAGAAAGAGG	397
Db	476	TAGTTTTTTAAGACAGCCTTTTGCCCCGGTCGACCCTACGGCAAATACATCAAGAAAGAGG	417
Qy	398	TGGGCAAGATCCTGTGCACTGACTGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCC	457
Db	416	TGGGCAAGATCCTGTGCACTGACTGCGCCACCCGGCCCAAGCTGAAGAAGATGAAGAGCC	357
Qy	458	AGACGGGACAGGTGGGTGAGAAGCAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCC	517
Db	356	AGACAGGAGAGGTGGGTGAGAAGCAGTCGCTCAAGTGTGAGGCAGCGGCGGGAAACCCCC	297

Qy 518 AGCCTTCCTACCGTTGGTTCAAGGATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCA 577
 |||| |||| || |||||||||||||||||| |||||| || || || |||||
 Db 296 AGCCCTCCTATCGCTGGTTCAAGGATGGCAAGGAACTCAACCGGAGTCGTGATATTCGCA 237
 Qy 578 TCAAATATGGCAACGGCAGAAAGAACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGG 637
 |||| |||||| |||||||||||||||||| |||||||||| |||| |||||
 Db 236 TCAAGTATGGCAATGGCAGAAAGAACTCACGGCTACAGTTCAACAAAGTGAGGGTGGAGG 177
 Qy 638 ACGCTGGGGAGTATGTCTGCGAGGCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCC 697
 | || |||||| |||| |||||||||||||| |||||||||| |||||
 Db 176 ATGCCGGGGAGTACGTCTGTGAGGCCGAGAACATCCTTGGGAAGGACACCGTGAAGGGGCC 117
 Qy 698 GGCTTTACGTCAACAGCGTGAGCACCACCCTGTCTATCCTGGTCGGGGCACGCCCCGGAAGT 757
 | || | |||||||||||||| |||||||||| || |||||||
 Db 116 GACTCCATGTCAACAGCGTGAGCACCCTGTCTATCCTGGTCGGGACATGCCCGGAAGT 57
 Qy 758 GCAACGAGACAGCCAAGTCCTA--TTGCGTCAATGGAGGCGTCTGCTACT 805
 |||| |||| |||||| || || || |||||| |||||
 Db 56 GCAATGAGACCGCCAAGTCCTACCATGTGTGAATGGAGGCGTGTGCTACT 7

RESULT 4

BX281777

LOCUS BX281777 524 bp mRNA linear EST 04-MAR-2003

DEFINITION BX281777 NIH_MGC_121 Homo sapiens cDNA clone IMAGp998K1811607 ;
 IMAGE:5240969, mRNA sequence.

ACCESSION BX281777

VERSION BX281777.1 GI:28612804

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 524)

AUTHORS Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,
 Radelof,U., Schneider,D. and Korn,B.

TITLE Human UnigeneSet - RZPD3

JOURNAL Unpublished

COMMENT Contact: Ina Rolfs

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
 RZPD; IMAGp998K1811607.

RZPDLIB; I.M.A.G.E. cDNA Clone Collection;

Human UnigeneSet - RZPD3 (RZPDLIB No.972)

[http://www.rzpd.de/CloneCards/cgi-](http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=972)

[bin/showLib.pl.cgi/response?libNo=972](http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=972) Contact: Ina Rolfs

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

Heubnerweg 6, D-14059 Berlin, Germany

Tel: +49 30 32639 101

Fax: +49 30 32639 111

www.rzpd.de

This clone is available royalty-free from RZPD;

contact RZPD (clone@rzpd.de) for further information. Seq primer:

M13u, Primer sequence: CGTTGTAAAACGACGGCCAGT.

FEATURES

Location/Qualifiers

source 1..524


```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGp998K1811607 ; IMAGE:5240969"
/lab_host="DH10B"
/clone_lib="NIH_MGC_121"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 3
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 017. Note:
this is a NIH_MGC Library."

```

```

BASE COUNT      99 a      174 c      172 g      79 t
ORIGIN

```

```

Query Match          47.0%; Score 467; DB 13; Length 524;
Best Local Similarity 100.0%; Pred. No. 5.5e-105;
Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCGCTCGCCTGC 60
        |||
Db      58 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCGCTCGCCTGC 117

Qy      61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
        |||
Db     118 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 177

Qy     121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
        |||
Db     178 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 237

Qy     181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
        |||
Db     238 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 297

Qy     241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
        |||
Db     298 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 357

Qy     301 CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
        |||
Db     358 CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 417

Qy     361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
        |||
Db     418 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 477

Qy     421 TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACA 467
        |||
Db     478 TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACA 524

```

RESULT 5

AA706226
 LOCUS AA706226 549 bp mRNA linear EST 12-JAN-1999
 DEFINITION ah28a07.s1 Soares_parathyroid_tumor_NbHPA Homo sapiens cDNA clone
 1240116 3' similar to TR:P43328 P43328 NEU DIFFERENTIATION FACTOR
 NDF04 ;, mRNA sequence.
 ACCESSION AA706226
 VERSION AA706226.1 GI:2716144
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 549)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
 , Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Possible reversed clone: similarity on wrong strand

Possible reversed clone: polyT not found

Insert Length: 689 Std Error: 0.00

Seq primer: -40m13 fwd. ET from Amersham

High quality sequence stop: 451.

FEATURES Location/Qualifiers
 source 1. .549
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="1240116"
 /tissue_type="parathyroid tumor"
 /dev_stage="adult"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares_parathyroid_tumor_NbHPA"
 /note="Organ: parathyroid gland; Vector: pT7T3D (Pharmacia
) with a modified polylinker; Site_1: Not I; Site_2: Eco
 RI; 1st strand cDNA was primed with a Not I - oligo(dT)
 primer
 [5'-TGTTACCAATCTGAAGTGGGAGCGCCGCGACCAATTTTTTTTTTTTTTTTTT
 TTTT-3'], double-stranded cDNA was size selected, ligated
 to Eco RI adapters (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of a modified pT7T3
 vector (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by Bento
 Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid
 adenomas was kindly provided by Dr. Stephen Marx, National
 Institute of Diabetes and Digestive and Kidney Diseases,
 NIH."

BASE COUNT 137 a 163 c 156 g 92 t 1 others

ORIGIN

Query Match 40.6%; Score 404; DB 9; Length 549;
 Best Local Similarity 98.5%; Pred. No. 2.2e-89;
 Matches 407; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

Qy      424 GCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAA 483
          |||
Db      15  GCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAA 74

Qy      484 TCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGAT 543
          |||
Db      75  TCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGAT 134

Qy      544 GGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAAC 603
          |||
Db      135 GGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAAC 194

Qy      604 TCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGCC 663
          |||
Db      195 TCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGCC 254

Qy      664 GAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCACC 723
          |||
Db      255 GAGAACATCCTGGGGAAGGACACCGTCCGGAGGCCGGCTTTACGTCAACAGCGTGAGCACC 314

Qy      724 ACCCTGTCATCCTGGTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCTTATTC 783
          |||
Db      315 ACCCTGTCATCCTGGTCGGGGCACGCCCCGGAAGTGCAACGNGACAGCCAAGTCTTATTC 374

Qy      784 GTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAA 836
          |||
Db      375 GTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAA 427
  
```

RESULT 6

AI041451

LOCUS AI041451 412 bp mRNA linear EST 28-AUG-1998

DEFINITION ow36c02.s1 Soares_parathyroid_tumor_NbHPA Homo sapiens cDNA clone
 IMAGE:1648898 3' similar to TR:O14511 O14511 NTAK. ;, mRNA
 sequence.

ACCESSION AI041451

VERSION AI041451.1 GI:3280645

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 412)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
 , Ph.D.

cdNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
 Insert Length: 671 Std Error: 0.00
 Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 1.

FEATURES
 source Location/Qualifiers
 1. .412
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1648898"
 /tissue_type="parathyroid tumor"
 /dev_stage="adult"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares_parathyroid_tumor_NbHPA"
 /note="Organ: parathyroid gland; Vector: pT7T3D (Pharmacia)
) with a modified polylinker; Site_1: Not I; Site_2: Eco
 RI; 1st strand cDNA was primed with a Not I - oligo(dT)
 primer
 [5'-TGTTACCAATCTGAAGTGGGAGCGCCGACCAATTTTTTTTTTTTTTTTTTTT
 TTTT-3'], double-stranded cDNA was size selected, ligated
 to Eco RI adapters (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of a modified pT7T3
 vector (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by Bento
 Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid
 adenomas was kindly provided by Dr. Stephen Marx, National
 Institute of Diabetes and Digestive and Kidney Diseases,
 NIH."

BASE COUNT 112 a 108 c 126 g 65 t 1 others
 ORIGIN

Query Match 39.8%; Score 395.6; DB 9; Length 412;
 Best Local Similarity 97.6%; Pred. No. 2.4e-87;
 Matches 401; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 426 CACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAATC 485
 |||
 Db 1 CACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAATC 60
 QY 486 GCTGAAGTGTGAGGCAGCAGCCGGAATCCCCAGCCTTCCTACCGTTGGTTCAAGGATGG 545
 |||
 Db 61 GCTGAAGTGTGAGGCAGCAGCGATAAATCCCCAGCCTTCCTACCGTTGGTTCAAGGATGG 120
 QY 546 CAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAACTC 605
 |||
 Db 121 CAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAACTC 180
 QY 606 ACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGCCGA 665
 |||
 Db 181 ACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGCCGA 240

Qy 666 GAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCACCAC 725
 |||
 Db 241 GAACATCCTGGGGAAGGACACCGTACGAGGCCGGCTTTACGTCAACAGCGTGACGACCAC 300
 Qy 726 CCTGTCATCCTGGTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTGCGT 785
 |||
 Db 301 CCTGTCATCCTGGTCGGGGCACGCCCCGGAAGTGCAACGNGACAGCCAAGTCCTATTGCGT 360
 Qy 786 CAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAA 836
 |||
 Db 361 CAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAA 411

RESULT 7

BX529505

ID BX529505 standard; RNA; EST; 488 BP.

XX

AC BX529505;

XX

SV BX529505.1

XX

DT 27-MAY-2003 (Rel. 75, Created)

DT 27-MAY-2003 (Rel. 75, Last updated, Version 1)

XX

DE RZPD Mus musculus cDNA clone IMAGp998N017639 = IMAGE:3153984 5' EST.

XX

KW EST; expressed sequence tag.

XX

OS Mus musculus (house mouse)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;

OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

XX

RN [1]

RP 1-488

RA Heil O., Ebert L., Neubert P., Peters M., Radelof U., Schneider D.,

RA Korn B.;

RT ;

RL Submitted (28-MAY-2003) to the EMBL/GenBank/DDBJ databases.

RL RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH Im Neuenheimer

RL Feld 580, D-69120 Heidelberg, Germany

XX

CC RZPD; IMAGp998N017639.

CC RZPDLIB; I.M.A.G.E. cDNA Clone Collection;

CC Mouse UnigeneSet - RZPD2 (RZPDLIB No.981)

CC <http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=981>

CC Contact: Ina Rolfs

CC RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

CC Heubnerweg 6, D-14059 Berlin, Germany

CC Tel: +49 30 32639 101

CC Fax: +49 30 32639 111

CC www.rzpd.de

CC This clone is available royalty-free from RZPD;

CC contact RZPD (clone@rzpd.de) for further information.

CC Seq primer: SP6, Primer sequence: ATTTAGGTGACACTATAG

XX

FH Key Location/Qualifiers

FH

FT source 1. .488
 FT /db_xref="taxon:10090"
 FT /note="Cloned unidirectionally. Primer: Oligo dT. Average
 FT insert 2 kb. Library constructed by Life Technologies,
 FT catalog #12017-018. Investigators providing samples: Lothar
 FT Hennighausen/Chu-Xia Deng, NIH Reference for transgenic
 FT model: Xu et al., Nature Genetics 22, 37-43 (1999). Note:
 FT this is a NCI_CGAP Library
 FT <http://www.ncbi.nlm.nih.gov/ncicgap/>."
 FT /organism="Mus musculus"
 FT /clone="IMAGp998N017639"
 FT /clone_lib="NCI_CGAP_Mam3 mammary tumor"
 FT /lab_host="DH10B"
 XX

SQ Sequence 488 BP; 129 A; 116 C; 149 G; 94 T; 0 other;

Query Match, 38.3%; Score 381; DB 4; Length 488;
 Best Local Similarity 91.0%; Pred. No. 1e-83;
 Matches 405; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Qy	469	GTGGGTGAGAAGCAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTAC	528
Db	1	GTGGGTGAGAAGCAGTCGCTCAAGTGTGAGGCAGCGGCGGAAACCCAGCCTTCCTAT	60
Qy	529	CGTTGGTTCAAGGATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGC	588
Db	61	CGCTGGTTCAAGGATGGCAAGGAACCTCAACCGGAGTCGTGATATTCGCATCAAGTATGGC	120
Qy	589	AACGGCAGAAAGAACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAG	648
Db	121	AATGGCAGAAAGAACTCACGGCTACAGTTCAACAAAGTGAGGGTGGAGGATGCCGGGGAG	180
Qy	649	TATGTCTGCGAGGCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTC	708
Db	181	TACGTCTGTGAGGCCGAGAACATCCTTGGGAAGGACACCGTGAGGGGCCGACTCCATGTC	240
Qy	709	AACAGCGTGAGCACCACCCTGTCTATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACA	768
Db	241	AACAGCGTGAGCACCCTGTCTATCCTGGTCGGGACATGCCCGGAAGTGCAATGAGACC	300
Qy	769	GCCAAGTCCTATTGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTC	828
Db	301	GCCAAGTCCTACTGTGTGAATGGAGGCGTGTGCTACTACATCGAGGGCATCAACCAGCTC	360
Qy	829	TCCTGCAAATGTCCAAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGA	888
Db	361	TCCTGCAAATGTCCAAACGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGA	420
Qy	889	TTGTACATGCCAGATCCTAAGCAAA	913
Db	421	TTGTACATGCCAGATCCTAAGCAAA	445

RESULT 8

BF108794

LOCUS BF108794 427 bp mRNA linear EST 20-OCT-2000
 DEFINITION 7152g03.x1 Soares_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone

Db 65 AGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 124

Qy 483 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 542
 |||

Db 125 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 184

Qy 543 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 602
 |||

Db 185 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 244

Qy 603 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 662
 |||

Db 245 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 304

Qy 663 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722
 |||

Db 305 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 364

Qy 723 CACCCTGTCATCCTGGTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 782
 |||

Db 365 CACCCTGTCATCCTGGTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 424

Qy 783 CGT 785
 |||

Db 425 CGT 427

RESULT 9

BI410828/c

LOCUS BI410828 949 bp mRNA linear EST 14-AUG-2001

DEFINITION 602963734F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5119065 5', mRNA sequence.

ACCESSION BI410828

VERSION BI410828.1 GI:15171751

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 949)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM11290 row: d column: 10

High quality sequence start: 28

High quality sequence stop: 919.

FEATURES

Location/Qualifiers

source

1. .949
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:5119065"
/tissue_type="pooled lung tumors"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NCI CGAP Lu33"
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA was prepared from mRNA obtained from pooled lung tumors with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCCTCTGTTTTTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 171 a 261 c 269 g 248 t
ORIGIN

Query Match 31.9%; Score 317.4; DB 12; Length 949;
Best Local Similarity 84.2%; Pred. No. 7.6e-68;
Matches 442; Conservative 0; Mismatches 71; Indels 12; Gaps 7;

Qy	397	GTGGGCAAGATCCTGTGCACTGACTGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGC	456
Db	947	GTGGGCCAGATCCTGGGCACTG-CTGCGCCACCCGCCCA-CTGAAGAAGATGAAGA-C	891
Qy	457	CAGACGGGACAGGTGGGTGAGAAGCAATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCC	516
Db	890	CAAACCAGAAGAGTCGGTGAGAACAGTTCGCTCAAGTGTGAGGCACGGCCGGGAAACCC	831
Qy	517	CAGCCTTCTACC-----GTTGGTTCAAGGATGGC-AAGGAGCTCAACCGCAGCCGAGAC	570
Db	830	CCCCACCCCTCCCTATCGCTGGTTTCAAGGATGGCAAAGGA-CTCAACCGGAGTCTGTGAT	771
Qy	571	ATTCGCATCAAATATGGCAACGGCAGAAAGAACTCACGACTACAGTTCAACAAGGTGAAG	630
Db	770	ATTCGCATCAAGTATGCCAATGGCAGAAAGAACTCACGGCTACAGTTCAACAAAAGTGAG	711
Qy	631	GT--GGAGGACGCTGGGGAGTATGTCTGCGAGGCCGAGAACATCCTGGGGAAGGACACCG	688
Db	710	GTTGGAGGATTGCCGGGGAGTACGTCTGTGAGGCCGAGAACATCCTTGGGAAGGACACCG	651
Qy	689	TCCGGGGCCGGCTTTACGTCAACAGCGTGAGCACCACCCTGTCATCCTGGTCGGGGCAGC	748
Db	650	TGA-GGGCCGACTCCATGTCAACAGCGTGAGCACCCTCTGTATCCTGGTCGGGACATG	592
Qy	749	CCCGGAAGTGCAACGAGACAGCCAAGTCTATTGCGTCAATGGAGGCGTCTGCTACTACA	808
Db	591	CCCGGAAGTGCAATGAGACCGCCAAGTCTACTGTGTGAATGGAGGCGTGTGCTACTACA	532
Qy	809	TCGAGGGCATCAACCAGCTCTCCTGCAAATGTCCAAATGGATTCTTCGGACAGAGATGTT	868
Db	531	TCGAGGGCATCAACCAGCTCTCCTGCAAATGTCCAAACGGATTCTTCGGACAGAGATGTT	472

Qy 869 TGGAGAACTGCCTTTGCGATTGTACATGCCAGATCCTAAGCAAA 913
 |||
 Db 471 TGGAGAACTGCCTTTGCGATTGTACATGCCAGATCCTAAGCAAA 427

RESULT 10

BI651936

LOCUS BI651936 795 bp mRNA linear EST 12-SEP-2001

DEFINITION 603298677F1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5339251 5', mRNA sequence.

ACCESSION BI651936

VERSION BI651936.1 GI:15566172

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 795)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM11861 row: j column: 20

High quality sequence stop: 795.

FEATURES

source

Location/Qualifiers

1..795

/organism="Mus musculus"

/mol_type="mRNA"

/strain="129,C57BL/6J,FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:5339251"

/tissue_type="tumor, gross tissue"

/dev_stage="10 months"

/lab_host="DH10B"

/clone_lib="NCI_CGAP_Mam3"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;

Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Library constructed by Life Technologies. Investigators

providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH

Reference for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999)."

BASE COUNT 204 a 226 c 219 g 146 t

ORIGIN

Query Match

27.4%; Score 272.6; DB 12; Length 795;

Best Local Similarity 92.3%; Pred. No. 8.8e-57;

Matches 298; Conservative 0; Mismatches 24; Indels 1; Gaps 1;

LOCUS BE983573 333 bp mRNA linear EST 29-APR-2002
DEFINITION UI-M-CG0p-bgi-c-07-0-UI.s1 NIH_BMAP_Ret4_S2 Mus musculus cDNA clone
UI-M-CG0p-bgi-c-07-0-UI 3', mRNA sequence.
ACCESSION BE983573
VERSION BE983573.1 GI:10654893
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 333)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mEST@mail.nih.gov
Oligo-dT track not found, Not I site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Soares Lab Clone distribution: Researchers may obtain BMAP cDNA
clones from RESEARCH GENETICS. It should be noted that Bento Soares
is generating a small number of additional specialized
non-redundant arrays of BMAP cDNAs whose availability will be

considered under appropriate and limited collaborative arrangements
The tissue for this library was contributed by Dr. Xin-Yuan Fu,
Yale University School of Medicine The following repetitive
elements were found in this cDNA sequence: 15-105,
>GC_rich#Low_complexity
Seq primer: M13 Forward
POLYA=No.

FEATURES Location/Qualifiers
source 1. .333
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-M-CG0p-bgi-c-07-0-UI"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NIH_BMAP_Ret4_S2"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The
NIH_BMAP_Ret4_S2 library is a subtracted library,
ultimately derived from mouse retina tissue libraries at
various stages of development. For a detailed description
of the library from which this clone was derived, please
visit our web site at brainest.eng.uiowa.edu. The tissue
for this library was contributed by Dr. Xin-Yuan Fu, Yale
University School of Medicine
TAG_SEQ=None found"

BASE COUNT 47 a 124 c 122 g 40 t
ORIGIN

Query Match 23.6%; Score 234.4; DB 10; Length 333;
Best Local Similarity 95.6%; Pred. No. 1.8e-47;
Matches 241; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC 60
 |||
Db 82 ATGAGGCGCGACCCGGCCCCCGGCTTCTCGATGCTGCTCTTCGGTGTGTCACTCGCCTGC 141
Qy 61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
 |||
Db 142 TACTCGCCCAGCCTCAAGTCGGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 201
Qy 121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
 |||
Db 202 GGCAAGGTACAGGGACTGGCCCCGGCAGGCGGTTCCAGCTCTAACAGCACCCGAGAGCCG 261
Qy 181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
 |||
Db 262 CCCGCCTCGGGTCGGGTGGCGCTGGTGAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 321
Qy 241 GGGCTGCAGCGC 252
 |||
Db 322 GGGCTGCAGCGC 333

RESULT 12

AW476657

LOCUS AW476657 529 bp mRNA linear EST 24-FEB-2000

DEFINITION uq79e01.y1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:2937336 5' similar to TR:O35073 O35073 NTAK ALPHA2-1P ;, mRNA sequence.

ACCESSION AW476657

VERSION AW476657.1 GI:7046763

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 529)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished

COMMENT Other_ESTs: uq79e01.x1
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

MGI:1049756
Seq primer: -40RP from Gibco
High quality sequence stop: 459.

FEATURES

source	Location/Qualifiers
	1. .529
	/organism="Mus musculus"
	/mol_type="mRNA"
	/strain="CZECH II"
	/db_xref="taxon:10090"
	/clone="IMAGE:2937336"
	/tissue_type="pooled lung tumors"
	/lab_host="DH10B (phage-resistant)"
	/clone_lib="NCI_CGAP_Lu33"
	/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA was prepared from mRNA obtained from pooled lung tumors with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCCTCTGTTTTTTTTTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 136 a 144 c 146 g 103 t

ORIGIN

Query Match 23.4%; Score 232.2; DB 9; Length 529;
Best Local Similarity 93.1%; Pred. No. 7.5e-47;
Matches 243; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 653 TCTGCGAGGCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACA 712

Db	2	TCTGTGAGGCCGAGAACATCCTTGGAAGGACACCGTGAGGGGCCGACTCCATGTCAACA	61
Qy	713	GCGTGAGCACCACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCA	772
Db	62	GCGTGAGCACCACCTCTGTCATCCTGGTCGGGACATGCCCGGAAGTGCAATGAGACCGCCA	121
Qy	773	AGTCCTATTGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCT	832
Db	122	AGTCCTACTGTGTGAATGGAGGCGTGTGCTACTACATCGAGGGCATCAACCAGCTCTCCT	181
Qy	833	GCAAATGTCCAAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGT	892
Db	182	GCAAATGTCCAAACGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGT	241
Qy	893	ACATGCCAGATCCTAAGCAAA	913
Db	242	ACATGCCAGATCCTAAGCAAA	262

AA772412

LOCUS	AA772412	297 bp	mRNA	linear	EST 31-DEC-1998
-------	----------	--------	------	--------	-----------------

DEFINITION ai44e12.s1 Soares_parathyroid_tumor_NbHPA Homo sapiens cDNA clone 135986 3' similar to TR:P43328 P43328 NEU DIFFERENTIATION FACTOR NDF04 ;, mRNA sequence.

ACCESSION AA772412

VERSION AA772412.1 GI:2824195

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 297)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

cdNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Possible reversed clone: similarity on wrong strand

Possible reversed clone: polyT not found

Insert Length: 667 Std Error: 0.00

Seq primer: -40ml3 fwd. ET from Amersham

High quality sequence stop: 267.

FEATURES

source

Location/Qualifiers

1. .297

```
/organism="Homo sapiens"
```

```
/mol type="mRNA"
```

```

/db_xref="taxon:9606"
/clone="1359886"
/tissue_type="parathyroid tumor"
/dev_stage="adult"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares_parathyroid_tumor_NbHPA"
/note="Organ: parathyroid gland; Vector: pT7T3D (Pharmacia
) with a modified polylinker; Site_1: Not I; Site_2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo(dT)
primer
[5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCACCAATTTTTTTTTTTTTTTTTT
TTTTT-3'], double-stranded cDNA was size selected, ligated
to Eco RI adapters (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of a modified pT7T3
vector (Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid
adenomas was kindly provided by Dr. Stephen Marx, National
Institute of Diabetes and Digestive and Kidney Diseases,
NIH."

```

```

BASE COUNT      87 a      68 c      105 g      37 t
ORIGIN

```

```

Query Match          22.5%; Score 224; DB 9; Length 297;
Best Local Similarity 92.2%; Pred. No. 6.5e-45;
Matches 236; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

```

```

Qy      405 GATCCTGTGCACTGACTGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGG 464
          || || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      42  GAGCCCGATCCCGGGAGAAAGCCACCCGGCCAAGTTGAAGAAGATGAAGAGCCAGACGGG 101

Qy      465 ACAGGTGGGTGAGAAGCAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTC 524
          || || || || || || || || || || || || || || || || || || || || ||
Db      102 ACAGGTGGGTGAGAAGCAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTC 161

Qy      525 CTACCGTTGGTTCAAGGATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATA 584
          || || || || || || || || || || || || || || || || || || || || ||
Db      162 CTACCGTTGGTTCAAGGATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATA 221

Qy      585 TGGCAACGGCAGAAAGAACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGG 644
          || || || || || || || || || || || || || || || || || || || || ||
Db      222 TGGCAACGGCAGAAAGAACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGG 281

Qy      645 GGAGTATGTCTGCGAG 660
          || || || || || || || || || || || || || || || || || || || || ||
Db      282 GGAGTATGTCTGCGAG 297

```

RESULT 14
BX089049/c

```

LOCUS      BX089049          362 bp      mRNA      linear      EST 23-JAN-2003
DEFINITION BX089049 Soares_parathyroid_tumor_NbHPA Homo sapiens cDNA clone
            IMAGp998M133119 ; IMAGE:1240116, mRNA sequence.
ACCESSION  BX089049
VERSION    BX089049.1  GI:27825909
KEYWORDS   EST.
SOURCE     Homo sapiens (human)

```

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 362)

AUTHORS Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,
 Radelof,U., Schneider,D. and Korn,B.

TITLE Human UnigeneSet - RZPD3

JOURNAL Unpublished

COMMENT Contact: Ina Rolfs
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
 RZPD; IMAGp998M133119.
 RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
 Human UnigeneSet - RZPD3 (RZPDLIB No.972)
<http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=972> Contact: Ina Rolfs
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Heubnerweg 6, D-14059 Berlin, Germany
 Tel: +49 30 32639 101
 Fax: +49 30 32639 111
www.rzpd.de
 This clone is available royalty-free from RZPD;
 contact RZPD (clone@rzpd.de) for further information. Seq primer:
 M13r, Primer sequence: TTTCACACAGGAAACAGCTATGAC.

FEATURES Location/Qualifiers

source 1. .362
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGp998M133119 ; IMAGE:1240116"
 /tissue_type="parathyroid tumor"
 /dev_stage="adult"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares_parathyroid_tumor_NbHPA"
 /note="Organ: parathyroid gland; Vector: pT7T3D (Pharmacia)
) with a modified polylinker; Site_1: Not I; Site_2: Eco
 RI; 1st strand cDNA was primed with a Not I - oligo(dT)
 primer
 [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGACCAATTTTTTTTTTTTTTTTTTTT
 TTTT-3'], double-stranded cDNA was size selected, ligated
 to Eco RI adapters (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of a modified pT7T3
 vector (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by Bento
 Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid
 adenomas was kindly provided by Dr. Stephen Marx, National
 Institute of Diabetes and Digestive and Kidney Diseases,
 NIH."

BASE COUNT 64 a 100 c 120 g 77 t 1 others

ORIGIN

Query Match 18.1%; Score 180; DB 13; Length 362;
 Best Local Similarity 99.4%; Pred. No. 5.5e-34;
 Matches 180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 656 GCGAGGCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCG 715
 |||

Db 362 GCGAGGCCGAGAACATCCTGGGGAAGGACACCGTCCGGGNCCGGCTTTACGTCAACAGCG 303
 QY 716 TGAGCACCACCCTGTCATCCTGGTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGT 775
 |||||
 Db 302 TGAGCACCACCCTGTCATCCTGGTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGT 243
 QY 776 CCTATTGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCA 835
 |||||
 Db 242 CCTATTGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCA 183
 QY 836 A 836
 |
 Db 182 A 182

RESULT 15

AW762061

LOCUS AW762061 256 bp mRNA linear EST 04-MAY-2000

DEFINITION ur53c01.y1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:3153984 5' similar to TR:O35073 O35073 NTAK ALPHA2-1P ;, mRNA sequence.

ACCESSION AW762061

VERSION AW762061.1 GI:7693978

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 256)

REFERENCE AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished

COMMENT Other_ESTs: ur53c01.x1

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

image.llnl.gov/image/html/iresources.shtml

MGI:1056740

Seq primer: -40RP from Gibco

High quality sequence stop: 161.

FEATURES

Location/Qualifiers

source

1. .256

/organism="Mus musculus"

/mol_type="mRNA"

/strain="129,C57BL/6J,FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:3153984"

/tissue_type="tumor, gross tissue"

/dev_stage="10 months"

/lab_host="DH10B"

/clone_lib="NCI_CGAP_Mam3"

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2004, 07:12:21 ; Search time 3976 Seconds
(without alignments)
10227.407 Million cell updates/sec

Title: US-09-864-675-1
Perfect score: 994
Sequence: 1 atgaggcgcgacccggcccc.....caccttggatttgaattaaa 994

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*

28: em_un:*
 29: em_vi:*
 30: em_htg_hum:*
 31: em_htg_inv:*
 32: em_htg_other:*
 33: em_htg_mus:*
 34: em_htg_pln:*
 35: em_htg_rod:*
 36: em_htg_mam:*
 37: em_htg_vrt:*
 38: em_sy:*
 39: em_htgo_hum:*
 40: em_htgo_mus:*
 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8							
Result	Query						
No.	Score	Match	Length	DB	ID	Description	
1	913.2	91.9	3020	9	AB005060	AB005060 Homo sapi	
2	900	90.5	1884	6	AR098145	AR098145 Sequence	
3	900	90.5	1884	6	AR116617	AR116617 Sequence	
4	881	88.6	993	6	AR072053	AR072053 Sequence	
5	800	80.5	2947	10	D89995	D89995 Rattus sp.	
6	799.4	80.4	3076	6	E16456	E16456 Rat mRNA fo	
7	799.4	80.4	3077	10	D89996	D89996 Rattus sp.	
8	738.6	74.3	3441	6	AR072052	AR072052 Sequence	
9	547.2	55.1	1607	6	AR098144	AR098144 Sequence	
10	547.2	55.1	1607	6	AR116616	AR116616 Sequence	
11	492	49.5	1476	6	AR098146	AR098146 Sequence	
12	492	49.5	1476	6	AR116618	AR116618 Sequence	
13	492	49.5	2268	6	AR098155	AR098155 Sequence	
14	492	49.5	2268	6	AR116627	AR116627 Sequence	
15	487.4	49.0	2188	10	AB001576	AB001576 Rattus sp	
16	467.8	47.1	2467	6	AR098143	AR098143 Sequence	
17	467.8	47.1	2467	6	AR116615	AR116615 Sequence	
18	424.8	42.7	118504	9	AC094080	AC094080 Homo sapi	
c 19	424.8	42.7	152838	2	AC011589	AC011589 Homo sapi	
20	424.8	42.7	170797	9	AC011379	AC011379 Homo sapi	
21	424.8	42.7	210675	2	AC026272	AC026272 Homo sapi	
22	424	42.7	1054	6	AX406616	AX406616 Sequence	
23	424	42.7	1054	9	HS2NRG01	AF119151 Homo sapi	
24	387.2	39.0	140307	2	AC131191	AC131191 Mus muscu	
25	384	38.6	253462	2	AC096477	AC096477 Rattus no	
26	359.6	36.2	1207	6	AR072054	AR072054 Sequence	
27	227.2	22.9	240	10	AY227025	AY227025 Mus muscu	
28	173	17.4	419	6	AX406617	AX406617 Sequence	
29	173	17.4	419	9	HS2NRG02	AF119152 Homo sapi	
30	173	17.4	120236	9	AC008523	AC008523 Homo sapi	
c 31	173	17.4	189050	9	AC008667	AC008667 Homo sapi	
32	142.6	14.3	85703	2	AC020830	AC020830 Mus muscu	
c 33	142.6	14.3	191101	2	AC127350	AC127350 Mus muscu	

	34	139.4	14.0	226038	2	AC106592	AC106592	Rattus no
	35	139.4	14.0	273080	2	AC098540	AC098540	Rattus no
c	36	139.4	14.0	302176	2	AC096479	AC096479	Rattus no
	37	135.6	13.6	142	6	AR072064	AR072064	Sequence
	38	124.6	12.5	493	6	AX406618	AX406618	Sequence
	39	124.6	12.5	493	9	HS2NRG03	AF119153	Homo sapi
	40	122.4	12.3	350	6	AX406619	AX406619	Sequence
	41	122.4	12.3	350	9	HS2NRG04	AF119154	Homo sapi
	42	108.4	10.9	206683	2	BX323592	BX323592	Danio rer
	43	108.4	10.9	220700	2	BX005008	BX005008	Danio rer
c	44	108	10.9	85703	2	AC020830	AC020830	Mus muscu
	45	95.4	9.6	1140	6	AR022498	AR022498	Sequence

ALIGNMENTS

RESULT 1

AB005060

LOCUS AB005060 3020 bp mRNA linear PRI 14-NOV-1997

DEFINITION Homo sapiens mRNA for NTAK, complete cds.

ACCESSION AB005060

VERSION AB005060.1 GI:2626738

KEYWORDS NTAK.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (sites)

AUTHORS Higashiyama, S., Horikawa, M., Yamada, K., Ichino, N., Nakano, N., Nakagawa, T., Miyagawa, J., Matsushita, N., Nagatsu, T., Taniguchi, N. and Ishiguro, H.

TITLE A novel brain-derived member of the epidermal growth factor family that interacts with ErbB3 and ErbB4

JOURNAL J. Biochem. 122 (3), 675-680 (1997)

MEDLINE 98006324

PUBMED 9348101

REFERENCE 2 (bases 1 to 3020)

AUTHORS Ishiguro, H.

TITLE Direct Submission

JOURNAL Submitted (24-JUN-1997) Hiroshi Ishiguro, Fujita Health University, ICMS; 1-98, Kutsukake-cho, Toyoake, Aichi 470-11, Japan (E-mail: hishi@fujita-hu.ac.jp, Tel: 0562-93-9393, Fax: 0562-93-8831)

FEATURES Location/Qualifiers

source 1..3020
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /cell_line="SK-N-SH"
 /cell_type="neuroblastoma"

CDS 226..2778
 /codon_start=1
 /product="NTAK"
 /protein_id="BAA23417.1"
 /db_xref="GI:2626739"
 /translation="MRQVCCSALPPPPLEKGRCSYSDSSSSSSSSSSSSSSSSSSSSSG
 SSSRSSNNSSISRPAAPPEPRPQQQPQPRSPAARRAAARSRAAAAGGMRRDPAPGFS"

MLLFGVSLACYSPSLKSVQDQAYKAPVVVEGKVQGLVPAGGSSSNSTREPPASGRVAL
VKVLDKWPLRSGGLQREQVISVGSCVPLERNQRYIFFLEPTEQPLVFKTAFAPLDTNG
KNLKKEVGKILCTDCATRPKLKMKMSQTGQVGEKQSLKCEAAAGNPQPSYRWFKDGE
LNRSRDIRIKYGNRKNRSLQFNKVVEDAGEYVCEAENILGKDTVGRGLVNSVSTT
LSSWSGHARKCNETAKSYCVNGGVCYYIEGINQLSCKCPNGFFGQRCLEKLPLRLYMP
DPKQKAEELYQKRVLTITGICVALLVVGIVCVVAYCKTKKQKQMHNLRLQNMCPAHQ
NRSLANGPSPRLDPEEIQMADYISKNPATDHVIRRETETTFSGSHSCSPSHHCSTA
TPTSSHRHESHTWSLERSESLTSDSQSGIMLSSVGTSKCNSPACVEARARRAAAYNLE
ERRRATAPPYHDSVDSLSDSPHSERYVSALTTPARLSPVDFHYSLATQVPTFEITSPN
SAHAVSLPPAAPISYRLAEQQPLLRLHPAPPGPGPGPGPGPGADMQRSYDSYYPAA
GPGPRRGTCALGGSLSLPASPFRIPEDEYETTQECAPPPPPRPRARGASRRTSAGP
RRWRRSRLNGLAAQARAARDSLSLSGSGGGSASASDDDDADDADGALAAESTPFLGL
RGAHDALRSDSPPLCPAADSRTYYSLDSHSTRASSRHSRGPPPRAKQDSAPL"

polyA_site

3020

/note="39 A nucleotides"

BASE COUNT 615 a 1015 c 937 g 453 t
ORIGIN

Query Match 91.9%; Score 913.2; DB 9; Length 3020;
Best Local Similarity 99.7%; Pred. No. 4.8e-171;
Matches 915; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC	60
Db	502	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC	561
Qy	61	TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Db	562	TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	621
Qy	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Db	622	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	681
Qy	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Db	682	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	741
Qy	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	742	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	801
Qy	301	CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	802	CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	861
Qy	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	862	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	921
Qy	421	TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Db	922	TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	981
Qy	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	540
Db	982	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	1041

```

Qy      541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG 600
          |||
Db      1042 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG 1101

Qy      601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660
          |||
Db      1102 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 1161

Qy      661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720
          |||
Db      1162 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 1221

Qy      721 ACCACCCTGTCATCCTGGTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 780
          |||
Db      1222 ACCACCCTGTCATCCTGGTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 1281

Qy      781 TCGCTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT 840
          |||
Db      1282 TCGCTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT 1341

Qy      841 CCAAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCA 900
          |||
Db      1342 CCAAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCA 1401

Qy      901 GATCCTAAGCAAAGTGTC 918
          |||
Db      1402 GATCCTAAGCAAAAAGCC 1419

```

RESULT 2

AR098145

LOCUS AR098145 1884 bp DNA linear PAT 14-FEB-2001

DEFINITION Sequence 5 from patent US 6074841.

ACCESSION AR098145

VERSION AR098145.1 GI:12807402

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 1884)

AUTHORS Gearing, D.P. and Busfield, S.J.

TITLE Don-1 gene and polypeptides and uses therefor

JOURNAL Patent: US 6074841-A 5 13-JUN-2000;

FEATURES Location/Qualifiers

source 1. .1884
/organism="unknown"

BASE COUNT 426 a 607 c 560 g 291 t

ORIGIN

Query Match 90.5%; Score 900; DB 6; Length 1884;

Best Local Similarity 99.3%; Pred. No. 2e-168;

Matches 914; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```

Qy      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCGCTCGCCTGC 60
          |||
Db      218 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCGCTCGCCTGC 277

```

Qy	61	TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Db	278	TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	337
Qy	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Db	338	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	397
Qy	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Db	398	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	457
Qy	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	458	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	517
Qy	301	CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	518	CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	577
Qy	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	578	CCCCT-GATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGGC	636
Qy	421	TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Db	637	TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	696
Qy	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	540
Db	697	CAATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	756
Qy	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAAG	600
Db	757	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAAG	816
Qy	601	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Db	817	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	876
Qy	661	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Db	877	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	936
Qy	721	ACCACCCTGTCATCCTGGTCCGGGACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	780
Db	937	ACCACCCTGTCATCCTGGTCCGGGACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	996
Qy	781	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT	840
Db	997	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT	1056
Qy	841	CCAAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCA	900
Db	1057	CCAAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCA	1116

Qy 901 GATCCTAAGCAAAGTGTCT 920
|||||||
Db 1117 GATCCTAAGCAAAGCACCT 1136

RESULT 3

AR116617

LOCUS AR116617 1884 bp DNA linear PAT 16-MAY-2001

DEFINITION Sequence 5 from patent US 6133423.

ACCESSION AR116617

VERSION AR116617.1 GI:14096939

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 1884)

AUTHORS Gearing, D.P. and Busfield, S.J.

TITLE Don-1 gene and polypeptides and uses therefor

JOURNAL Patent: US 6133423-A 5 17-OCT-2000;

FEATURES Location/Qualifiers

source 1. 1884

/organism="unknown"

BASE COUNT 426 a 607 c 560 g 291 t

ORIGIN

Query Match 90.5%; Score 900; DB 6; Length 1884;

Best Local Similarity 99.3%; Pred. No. 2e-168;

Matches 914; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 1 ATGAGGCGGACCCGGCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCGCTCGCCTGC 60
|||||||
Db 218 ATGAGGCGGACCCGGCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCGCTCGCCTGC 277

Qy 61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
|||||||
Db 278 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 337

Qy 121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
|||||||
Db 338 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 397

Qy 181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
|||||||
Db 398 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 457

Qy 241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
|||||||
Db 458 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 517

Qy 301 CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
|||||||
Db 518 CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 577

Qy 361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
|||||||
Db 578 CCCCT-GATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGGC 636

Qy 421 TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480
 |||
 Db 637 TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 696
 Qy 481 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 540
 |||
 Db 697 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 756
 Qy 541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG 600
 |||
 Db 757 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG 816
 Qy 601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660
 |||
 Db 817 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 876
 Qy 661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720
 |||
 Db 877 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 936
 Qy 721 ACCACCCTGTCATCCTGGTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 780
 |||
 Db 937 ACCACCCTGTCATCCTGGTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 996
 Qy 781 TCGCTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT 840
 |||
 Db 997 TCGCTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT 1056
 Qy 841 CCAAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCA 900
 |||
 Db 1057 CCAAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCA 1116
 Qy 901 GATCCTAAGCAAAGTGTCTCT 920
 |||
 Db 1117 GATCCTAAGCAAAAGCACCT 1136

RESULT 4

AR072053

LOCUS AR072053 993 bp DNA linear PAT 18-FEB-2000

DEFINITION Sequence 3 from patent US 5912326.

ACCESSION AR072053

VERSION AR072053.1 GI:7222941

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 993)

AUTHORS Chang,H.

TITLE Cerebellum-derived growth factors

JOURNAL Patent: US 5912326-A 3 15-JUN-1999;

FEATURES Location/Qualifiers

source 1. .993

/organism="unknown"

BASE COUNT 230 a 271 c 311 g 181 t

ORIGIN

Query Match 88.6%; Score 881; DB 6; Length 993;
 Best Local Similarity 93.0%; Pred. No. 1.2e-164;
 Matches 923; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

Qy	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC	60
Db	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCGATGCTGCTCTTCGGTGTGTCACTCGCCTGC	60
Qy	61	TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Db	61	TACTCGCCCAGCCTCAAGTCCGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Qy	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Db	121	GGCAAGGTACAGGGACTGGCCCCGGCAGGCGGTTCCAGCTCTAACAGCACCCGAGAGCCT	180
Qy	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Db	181	CCCGCCTCGGGTCGGGTGGCGCTGGTGAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Qy	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGCGCGCCGCTCGAAAGGAACCAG	300
Qy	301	CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	301	CGCTACATCTTTTTCCTGGAGCCACCGAGCAGCCCTTAGTTTTTAAGACAGCCTTTGCC	360
Qy	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	361	CCGGTCGACCCTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Qy	421	TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Db	421	TGCGCAACCCGGCCCCAAGCTGAAGAAGATGAAGAGTCAGACAGGAGAGGTGGGCGAGAAG	480
Qy	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	540
Db	481	CAGTCGCTCAAGTGTGAGGCGGCGGCGGGGAACCCCAAGCCCTCCTATCGATGGTTCAAG	540
Qy	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTGCGCATCAAATATGGCAACGGCAGAAAG	600
Db	541	GACGGCAAGGAGCTCAACCGGAGTCGTGACATTGCGCATCAAGTATGGCAACGGCAGAAAG	600
Qy	601	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Db	601	AACTCACGGCTACAGTTCAACAAAGTGAAGGTGGAGGACGCTGGAGAGTACGTCTGTGAG	660
Qy	661	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Db	661	GCTGAGAACATCCTTGGGAAGGACACTGTGAGGGGCGGCTCCATGTCAACAGTGTGAGC	720
Qy	721	ACCACCCTGTCATCCTGGTCGGGGCACGCCCAGGAGTGAACGAGACAGCCAAGTCCTAT	780
Db	721	ACCACTCTGTCGTCTGGTCGGGGCACGCCCAGGAGTGAATGAGACAGCCAAGTCCTAC	780
Qy	781	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT	840

SRSSSRSSRGSTTTTSSSENSGNSGSI FRPAAPPEPRPQPQPRSPAARRAAARS
AAAAGGMRRDPAPGSSMLLFVSLACYSPSLKSVQDQAYKAPVVVEGKVQGLAPAGGS
SSNSTREPPASGRVALVKVLDKWPLRSGGLQREQVISVGSCAPLERNQRYIFFLEPTE
QPLVFKTAFAFVDPNGKNIKKEVGKILCTDCATRPKLKMKMSQTGEVGEKQSLKCEAA
AGNPQPSYRWFKD GKELNRSRDIRIKYGNRKN SRLQFNKVVEDAGEYVCEAENILG
KDTVGRGLHVNSVSTTLSSWSGHARKCNETAKSYCVNGGVCYYIEGINQLSCKCPNGF
FGQRCLEKLPLRLYMPDPKQKHLGFELKEAEELYQKRVLTITGICVALLVVGIVCVVA
YCKTKKQRRQMHHHLRQNMCPAHQNRSLANGPSHPRLDPEEIQMADYISKVNPATDHV
IRREAETTFSGSHSCSPSHHCSTATPTSSHRHESHTWSLERSESLTSDSQSGIMLSSV
GTSKCN SPACVEARARRAAAYSQEERRRAAMPYHDSIDSLRDSPHSERYVSALTTPA
RLSPVDFHYSLATQVPTFEITSPNSAHAVSLPPAAPISYRLAEQQPLLRHPAPPGP
GPGADMQRSYDSYYPAAGPGPRRGACALGGSLSLPASPFRIPEDEYETTQECAPP
PPPRPRTRGASRRTSAGPRRWRRSRLNGLAAQRARAARDSLSLSSGSGCGSASASDDD
ADDADGALAAESTPFLGLRAAHDALRSDSPPLCPAADSRTYYSLDSHSTRASSRHSRG
PPTRAKQDSGPL"

BASE COUNT 665 a 945 c 895 g 442 t
ORIGIN

Query Match 80.5%; Score 800; DB 10; Length 2947;
Best Local Similarity 91.8%; Pred. No. 1.4e-148;
Matches 845; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

Qy	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCGCTCGCCTGC	60
Db	403	ATGAGGCGCGACCCGGCCCCCGGCTCCTCGATGCTGCTCTTCGGTGTGTCCTCGCCTGC	462
Qy	61	TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Db	463	TACTCGCCCAGCCTCAAGTCCGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	522
Qy	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Db	523	GGCAAGGTACAGGGACTGGCCCCGGCAGGCGGTTCCAGCTCTAACAGCACCCGAGAGCCT	582
Qy	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Db	583	CCCGCCTCGGGTCGGGTGGCGCTGGTGAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	642
Qy	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	643	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGCGCGCCGCTCGAAAGGAACCAG	702
Qy	301	CGCTACATCTTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	703	CGCTACATCTTTTTCTGGAGCCACCGAGCAGCCCTTAGTTTTTAAGACAGCCTTTGCC	762
Qy	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	763	CCGTCGACCCTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	822
Qy	421	TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Db	823	TGCGCAACCCGGCCCAAGCTGAAGAAGATGAAGAGTCAGACAGGAGAGGTGGGCGAGAAG	882
Qy	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCTACCGTTGGTTCAAG	540
Db	883	CAGTCGCTCAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCTATCGATGGTTCAAG	942

Qy 541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG 600
 || |||||
 Db 943 GACGGCAAGGAGCTCAACCGGAGTCGTGACATTCGCATCAAGTATGGCAACGGCAGAAAG 1002
 Qy 601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660
 |||||
 Db 1003 AACTCACGGCTACAGTTCAACAAAGTGAAGGTGGAGGACGCTGGAGAGTACGTCTGTGAG 1062
 Qy 661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720
 || |||||
 Db 1063 GCTGAGAACATCCTTGGGGAAGGACACTGTGAGGGGCCGGCTCCATGTCAACAGTGTGAGC 1122
 Qy 721 ACCACCCTGTCATCCTGGTTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 780
 |||||
 Db 1123 ACCACTCTGTCGTCTGCTCGGGGCACGCCCGGAAGTGCAATGAGACAGCCAAGTCCTAC 1182
 Qy 781 TCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT 840
 || || |||||
 Db 1183 TGTGTGAATGGAGGCGTGTGCTACTACATCGAAGGCATCAACCAACTCTCCTGCAAATGT 1242
 Qy 841 CCAAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCA 900
 |||||
 Db 1243 CCAAACGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCA 1302
 Qy 901 GATCCTAAGCAAAGTGTCTCT 920
 |||||
 Db 1303 GATCCTAAGCAAAGCACCT 1322

RESULT 6

E16456

LOCUS E16456 3076 bp DNA linear PAT 28-JUL-1999

DEFINITION Rat mRNA for neuregulin-like Transmembrane Activator for ErbB Kinases (NTAK).

ACCESSION E16456

VERSION E16456.1 GI:5711139

KEYWORDS JP 1998179166-A/1.

SOURCE Rattus sp.

ORGANISM Rattus sp.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 3076)

AUTHORS Higashiyama,S., Taniguchi,N., Ishiguro,K. and Nagatsu,T.

TITLE GENE ENCODING RECEPTOR TYPE TYROSINE-KINASE ERB B LIGAND AND ITS

JOURNAL Patent: JP 1998179166-A 1 07-JUL-1998;

HIGASHIYAMA SHIGEKI

COMMENT OS Rattus sp. (rat)

PN JP 1998179166-A/1

PD 07-JUL-1998

PF 25-DEC-1996 JP 1996356998

PI HIGASHIYAMA SHIGEKI, TANIGUCHI NAOYUKI, ISHIGURO KEIJI, PI NAGATSU TOSHIHARU

PC C12N15/09,C07K14/705,C07K16/28,C12N5/10,C12N15/02,C12P21/02,

PC C12P21/08,

PC C12Q1/68,G01N33/53,G01N33/566//A61K31/70,A61K38/46,A61K39/395,

PC A61K48/00,
 PC C07H21/04, (C12N5/10, C12R1:91), (C12P21/02, C12R1:91); CC
 strandedness: Double;
 CC topology: Linear;
 FH Key Location/Qualifiers
 FH
 FT source 1. .3076
 FT /organism='Rattus sp.'
 FT /cell_line='PC12'
 FT CDS 232. .2814
 FT /product='NTAK protein'.

FEATURES Location/Qualifiers
 source 1. .3076
 /organism="Rattus sp."
 /mol_type="genomic DNA"
 /db_xref="taxon:10118"

BASE COUNT 673 a 996 c 944 g 463 t
 ORIGIN

Query Match 80.4%; Score 799.4; DB 6; Length 3076;
 Best Local Similarity 92.2%; Pred. No. 1.8e-148;
 Matches 842; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

Qy	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC	60
Db	556	ATGAGGCGCGACCCGGCCCCCGGCTCCTCGATGCTGCTCTTCGGTGTGTCACTCGCCTGC	615
Qy	61	TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Db	616	TACTCGCCCAGCCTCAAGTCCGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	675
Qy	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Db	676	GGCAAGGTACAGGGACTGGCCCCGGCAGGCGGTTCCAGCTCTAACAGCACCCGAGAGCCT	735
Qy	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Db	736	CCCGCCTCGGGTCGGGTGGCGCTGGTGAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	795
Qy	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	796	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGCGCGCCGCTCGAAAGGAACCAG	855
Qy	301	CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	856	CGCTACATCTTTTTCCTGGAGCCACCGAGCAGCCCTTAGTTTTTAAGACAGCCTTTGCC	915
Qy	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	916	CCGGTCGACCCTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	975
Qy	421	TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Db	976	TGCGCAACCCGGCCCCAAGCTGAAGAAGATGAAGAGTCAGACAGGAGAGGTGGGCGAGAAG	1035
Qy	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCTACCGTTGGTTCAAG	540

Db	1036	CAGTCGCTCAAGTGTGAGGCGGCGGCGGGGAACCCCCAGCCCTCCTATCGATGGTTCAAG	1095
Qy	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG	600
Db	1096	GACGGCAAGGAGCTCAACCGGAGTCGTGACATTTCGCATCAAGTATGGCAACGGCAGAAAG	1155
Qy	601	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Db	1156	AACTCACGGCTACAGTTCAACAAAGTGAAGGTGGAGGACGCTGGAGAGTACGTCTGTGAG	1215
Qy	661	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Db	1216	GCTGAGAACATCCTTGGGAAGGACACTGTGAGGGGCCGGCTCCATGTCAACAGTGTGAGC	1275
Qy	721	ACCACCCTGTTCATCCTGGTTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	780
Db	1276	ACCACTCTGTTCGTCTGGTTCGGGGCACGCCCGGAAGTGCAATGAGACAGCCAAGTCCTAC	1335
Qy	781	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT	840
Db	1336	TGTGTGAATGGAGGCGTGTGCTACTACATCGAAGGCATCAACCAACTCTCCTGCAAATGT	1395
Qy	841	CCAAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCA	900
Db	1396	CCAAACGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCA	1455
Qy	901	GATCCTAAGCAAA	913
Db	1456	GATCCTAAGCAAA	1468

JOURNAL Submitted (21-DEC-1996) Hiroshi Ishiguro, Fujita Health University,
ICMS; 1-98, kutsukake-cho, Toyoake, Aichi 470-11, Japan

(E-mail:hishi@fujita-hu.ac.jp, Tel:0562-93-9393, Fax:0562-93-8831)

FEATURES
source Location/Qualifiers
1. .3077
/organism="Rattus sp."
/mol_type="mRNA"
/db_xref="taxon:10118"
/cell_line="PC12"
/cell_type="pheochromocytoma"
CDS
233. .2815
/codon_start=1
/product="NTAK alpha2"
/protein_id="BAA23345.1"
/db_xref="GI:2605632"
/translation="MRQVCCSALPPPLEKARCSSYSYSDSSSSSSSSNNSSSSTSSRSS
SRSSSRSSRGSTTTTSSSENSGNSGSI FRPAAPPEPRPQPPQPRSPAARRAAARS
AAAAGGMRRDPAPGSSMLLFGVSLACYSPSLKSVQDQAYKAPVVEGKVQGLAPAGGS
SSNSTREPPASGRVALVKVLDKWPLRSGGLQREQVISVGSCAPLERNQRYIFFLEPTE
QPLVFKTAFAFVDPNGKNIKKEVGKILCTDCATRPKLKMKSQTEVGEKQSLKCEAA
AGNPQPSYRWFKDGKELNRSRDIRIKYGNRKN SRLQFNKVKVEDAGEYVCEAENILG
KDTVGRHLVNSVSTTLSSWSGHARKCNETAKSYCVNGGVCYYIEGINQLSCKCPNGF
FGQRCLEKLPLRLYMPDPKQKAEELYQKRVLTTITGICVALLVVGIVCVVAYCKTKKQR
RQMHHHLRQNMCPAHQNRSLANGPSHPRLDPEEIQMADYISKNPATDHVIRREAETT
FSGSHSCSPSHHCSTATPTSSHRHESHTWSLSESLTSDSQSGIMLSSVGTSKCNSP
ACVEARARRAAAYSQEERRRAAMPYPYHDSIDSLRDSPHSERYVSALTTPARLSPVDFH
YSLATQVPTFEITSPNSAHAVSLPPAAPISYRLAEQQPLL RHPAPPGP GPGADMQR
SYDSYYYPAAGPGPRRGACALGGS LGSLPASPFRI PEDDEYETTQECAPPPPRPRTR
GASRRTSAGPRRWRRLNGLAAQRARAARDLSLSLSSGSGCSASASDDADADGAL
AAESTPFLGLRAAHDALRSDSPPLCPAADSRTYYSLDSHSTRASSRHSRGPPTRAKQD
SGPL"
BASE COUNT 673 a 996 c 945 g 463 t
ORIGIN

Query Match 80.4%; Score 799.4; DB 10; Length 3077;
Best Local Similarity 92.2%; Pred. No. 1.8e-148;
Matches 842; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

Qy 1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC 60
|||||
Db 557 ATGAGGCGCGACCCGGCCCCCGGCTCCTCGATGCTGCTCTTCGGTGTGTCACTCGCCTGC 616
Qy 61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
|||||
Db 617 TACTCGCCCAGCCTCAAGTCCGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 676
Qy 121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
|||||
Db 677 GGCAAGGTACAGGGACTGGCCCCGGCAGGCGGTTCCAGCTCTAACAGCACCCGAGAGCCT 736
Qy 181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
|||||
Db 737 CCCGCCTCGGGTCGGGTGGCGCTGGTGAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 796
Qy 241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCGAG 300
|||||
Db 797 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGCGCGCCGCTCGAAAGGAACCGAG 856
Qy 301 CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360

```

      |||
Db      857 CGCTACATCTTTTTCCTGGAGCCCACCGAGCAGCCCTTAGTTTTAAAGACAGCCTTTGCC 916
      |||
Qy      361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
      |||
Db      917 CCGGTCGACCCTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 976
      |||
Qy      421 TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480
      |||
Db      977 TGCGCAACCCGGCCCCAAGCTGAAGAAGATGAAGAGTCAGACAGGAGAGGTGGGCGAGAAG 1036
      |||
Qy      481 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 540
      |||
Db      1037 CAGTCGCTCAAGTGTGAGGCGGCGGGGAACCCCCAGCCCTCCTATCGATGGTTCAAG 1096
      |||
Qy      541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG 600
      |||
Db      1097 GACGGCAAGGAGCTCAACCGGAGTCGTGACATTCGCATCAAGTATGGCAACGGCAGAAAG 1156
      |||
Qy      601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660
      |||
Db      1157 AACTCACGGCTACAGTTCAACAAAGTGAAGGTGGAGGACGCTGGAGAGTACGTCTGTGAG 1216
      |||
Qy      661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720
      |||
Db      1217 GCTGAGAACATCCTTGGGAAGGACACTGTGAGGGGCCGGCTCCATGTCAACAGTGTGAGC 1276
      |||
Qy      721 ACCACCCTGTCATCCTGGTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 780
      |||
Db      1277 ACCACTCTGTCGTCTGGTCGGGGCAGCCCCGGAAGTGCAATGAGACAGCCAAGTCCTAC 1336
      |||
Qy      781 TCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT 840
      |||
Db      1337 TGTGTGAATGGAGGCGTGTGCTACTACATCGAAGGCATCAACCAACTCTCCTGCAAATGT 1396
      |||
Qy      841 CCAAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCA 900
      |||
Db      1397 CCAACGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCA 1456
      |||
Qy      901 GATCCTAAGCAAA 913
      |||
Db      1457 GATCCTAAGCAAA 1469

```

RESULT 8

AR072052

LOCUS AR072052 3441 bp DNA linear PAT 18-FEB-2000

DEFINITION Sequence 1 from patent US 5912326.

ACCESSION AR072052

VERSION AR072052.1 GI:7222940

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 3441)

AUTHORS Chang, H.

TITLE Cerebellum-derived growth factors

JOURNAL Patent: US 5912326-A 1 15-JUN-1999;

FEATURES Location/Qualifiers

source 1. .3441

/organism="unknown"

BASE COUNT 777 a 1057 c 1015 g 592 t

ORIGIN

Query Match 74.3%; Score 738.6; DB 6; Length 3441;

Best Local Similarity 90.4%; Pred. No. 2.1e-136;

Matches 789; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

```
Qy      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCCTGC 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      180 ATGAGGCGCGACCCGGCCCCCGGCTTCTCGATGCTGCTCTTCGGTGTGTCACTCGCCTGC 239

Qy      61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      240 TACTCGCCCAGCCTCAAGTCCGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 299

Qy     121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
        |||||||||||||||| ||| ||| || ||||| ||||||| |||||||
Db     300 GGCAAGGTACAGGGACTGGCCCCGGCAGGCGGTTCCAGCTCTAACAGCACCCGAGAGCCT 359

Qy     181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
        |||||||||||||||||| |||| |||||||||||||||||||||||
Db     360 CCCGCCTCGGGTCGGGTGGCGCTGGTGAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 419

Qy     241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
        |||||||||||||||||||||||||||||||||||||||| | |||||||
Db     420 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGCGCGCCGCTCGAAAGGAACCAG 479

Qy     301 CGCTACATCTTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
        |||||||||||||||||| || |||||||||| ||||||| |||||||
Db     480 CGCTACATCTTTTTCTGGAGCCACCGAGCAGCCCTTAGTTTTTAAGACAGCCTTTGCC 539

Qy     361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
        || |||| | |||||||||| |||||||||||||||||||||||
Db     540 CCGGTCGACCCTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 599

Qy     421 TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480
        ||||| ||||||| ||||||| ||||| ||| ||||||| |||||||
Db     600 TGCGCAACCCGGCCCCAAGCTGAAGAAGATGAAGAGTCAGACAGGAGAGGTGGGCGAGAAG 659

Qy     481 CAATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCTACCGTTGGTTCAAG 540
        || ||||| ||||||| || || || || ||||| ||||| || |||||||
Db     660 CAGTCGCTCAAGTGTGAGGCGGCGGGGAACCCCGAGCCCTCTATCGATGGTTCAAG 719

Qy     541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG 600
        || |||||||||||||| || || |||||||||| ||||||| |||||||
Db     720 GACGGCAAGGAGCTCAACCGGAGTCGTGACATTCGCATCAAGTATGGCAACGGCAGAAAG 779

Qy     601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660
        ||||||| ||||||| ||||||| ||||||| ||||| |||
Db     780 AACTCACGGCTACAGTTCAACAAAGTGAAGGTGGAGGACGCTGGAGAGTACGTCTGTGAG 839

Qy     661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720
        || ||||||| ||||||| || ||||||| | ||||||| |||||
```

[illegible]

Db 62 GGCCCAAGCTGAAGAAGATGAAGAGCCAGACAGGAGAGGTGGGTGAGAAGCAGTCGCTCA 121
 Qy 491 AGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGATGGCAAGG 550
 |||||
 Db 122 AGTGTGAGGCAGCGGCGGGAAACCCAGCCCTCCTATCGCTGGTTCAAGGATGGCAAGG 181
 Qy 551 AGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAACTCACGAC 610
 |
 Db 182 AACTCAACCGGAGTCGTGATATTCGCATCAAGTATGGCAATGTCAGAAAGAACTCACGGC 241
 Qy 611 TACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGCCGAGAACA 670
 |||||
 Db 242 TACAGTTCAACAAAGTGAGGGTGGAGGATGCCGGGGAGTACGTCTGTGAGGCCGAGAACA 301
 Qy 671 TCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCACCACCCTGT 730
 ||||
 Db 302 TCCTTGGGAAGGACACCGTGAGGGGCCGACTCCATGTCAACAGCGTGAGCACCCTGT 361
 Qy 731 CATCCTGGTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTGCGTCAATG 790
 |||||
 Db 362 CATCCTGGTCGGGACATGCCCGGAAGTGCAATGAGACCGCCAAGTCCTACTGTGTGAATG 421
 Qy 791 GAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCCAAATGGAT 850
 |||||
 Db 422 GAGGCGTGTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCCAAACGGAT 481
 Qy 851 TCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCAGATCCTAAGC 910
 |||||
 Db 482 TCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCAGATCCTAAGC 541
 Qy 911 AAAGTGTCTGTGGGATACACCGGGGACAGGTGTCAGCAGTTCGCAATGGTCAACTTCTC 970
 |||||
 Db 542 AAAGTGTCTGTGGGATACACCGGGGACAGGTGTCAGCAGTTCGCAATGGTCAACTTCTC 601
 Qy 971 CAAGCACCTTGGATTTGAATTAAA 994
 |||||
 Db 602 CAAGCACCTTGGATTTGAATTGAA 625

RESULT 11

AR098146

LOCUS AR098146 1476 bp DNA linear PAT 14-FEB-2001

DEFINITION Sequence 7 from patent US 6074841.

ACCESSION AR098146

VERSION AR098146.1 GI:12807403

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 1476)

AUTHORS Gearing,D.P. and Busfield,S.J.

TITLE Don-1 gene and polypeptides and uses therefor

JOURNAL Patent: US 6074841-A 7 13-JUN-2000;

FEATURES Location/Qualifiers

source 1. .1476

/organism="unknown"

BASE COUNT 335 a 473 c 452 g 216 t

ORIGIN

Query Match 49.5%; Score 492; DB 6; Length 1476;
 Best Local Similarity 92.8%; Pred. No. 1.7e-87;
 Matches 516; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

```

Qy      363 CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG 422
          || || | || | | | | | | | | | | | | | |
Db      98 CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCCGATCCCGGGGAGAA 157

Qy      423 CGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 482
          |||||
Db      158 AGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 217

Qy      483 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 542
          |||||
Db      218 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 277

Qy      543 TGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAA 602
          |||||
Db      278 TGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAA 337

Qy      603 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 662
          |||||
Db      338 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 397

Qy      663 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722
          |||||
Db      398 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 457

Qy      723 CACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 782
          |||||
Db      458 CACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 517

Qy      783 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC 842
          |||||
Db      518 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC 577

Qy      843 AAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCAGA 902
          |||||
Db      578 AAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCAGA 637

Qy      903 TCCTAAGCAAAGTGTC 918
          |||||
Db      638 TCCTAAGCAAAAAGCC 653
  
```

RESULT 12

AR116618

LOCUS AR116618 1476 bp DNA linear PAT 16-MAY-2001
 DEFINITION Sequence 7 from patent US 6133423.
 ACCESSION AR116618
 VERSION AR116618.1 GI:14096940
 KEYWORDS .
 SOURCE Unknown.
 ORGANISM Unknown.
 Unclassified.

REFERENCE 1 (bases 1 to 1476)
 AUTHORS Gearing,D.P. and Busfield,S.J.
 TITLE Don-1 gene and polypeptides and uses therefor
 JOURNAL Patent: US 6133423-A 7 17-OCT-2000;
 FEATURES Location/Qualifiers
 source 1. .1476
 /organism="unknown"
 BASE COUNT 335 a 473 c 452 g 216 t
 ORIGIN

Query Match 49.5%; Score 492; DB 6; Length 1476;
 Best Local Similarity 92.8%; Pred. No. 1.7e-87;
 Matches 516; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Qy	363	CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG	422
Db	98	CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCGATCCCGGGGAGAA	157
Qy	423	CGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA	482
Db	158	AGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA	217
Qy	483	ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA	542
Db	218	ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA	277
Qy	543	TGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAA	602
Db	278	TGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAA	337
Qy	603	CTCAGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC	662
Db	338	CTCAGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC	397
Qy	663	CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC	722
Db	398	CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC	457
Qy	723	CACCCGTGTCATCCTGGTTCGGGGCACGCCCAGGAGTGAACGAGACAGCCAAGTCCTATTG	782
Db	458	CACCCGTGTCATCCTGGTTCGGGGCACGCCCAGGAGTGAACGAGACAGCCAAGTCCTATTG	517
Qy	783	CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC	842
Db	518	CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC	577
Qy	843	AAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCAGA	902
Db	578	AAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCAGA	637
Qy	903	TCCTAAGCAAAGTGTC	918
Db	638	TCCTAAGCAAAAAGCC	653

RESULT 13
 AR098155

LOCUS AR098155 2268 bp DNA linear PAT 14-FEB-2001
 DEFINITION Sequence 31 from patent US 6074841.
 ACCESSION AR098155
 VERSION AR098155.1 GI:12807412
 KEYWORDS .
 SOURCE Unknown.
 ORGANISM Unknown.
 Unclassified.
 REFERENCE 1 (bases 1 to 2268)
 AUTHORS Gearing, D.P. and Busfield, S.J.
 TITLE Don-1 gene and polypeptides and uses therefor
 JOURNAL Patent: US 6074841-A 31 13-JUN-2000;
 FEATURES Location/Qualifiers
 source 1..2268
 /organism="unknown"
 BASE COUNT 502 a 734 c 701 g 331 t
 ORIGIN

Query Match 49.5%; Score 492; DB 6; Length 2268;
 Best Local Similarity 92.8%; Pred. No. 1.7e-87;
 Matches 516; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Qy	363	CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG	422
Db	98	CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCGATCCCGGGGAGAA	157
Qy	423	CGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA	482
Db	158	AGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA	217
Qy	483	ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA	542
Db	218	ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA	277
Qy	543	TGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAA	602
Db	278	TGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAA	337
Qy	603	CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC	662
Db	338	CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC	397
Qy	663	CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC	722
Db	398	CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC	457
Qy	723	CACCCTGTCATCCTGGTCGGGGCACGCCCAGGAGTGAACGAGACAGCCAAGTCCTATTG	782
Db	458	CACCCTGTCATCCTGGTCGGGGCACGCCCAGGAGTGAACGAGACAGCCAAGTCCTATTG	517
Qy	783	CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC	842
Db	518	CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC	577
Qy	843	AAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCAGA	902
Db	578	AAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCAGA	637

Qy 903 TCCTAAGCAAAGTGTC 918
| | | | | | | | | | | |
Db 638 TCCTAAGCAAAAAGCC 653

RESULT 14

AR116627

LOCUS AR116627 2268 bp DNA linear PAT 16-MAY-2001

DEFINITION Sequence 31 from patent US 6133423.

ACCESSION AR116627

VERSION AR116627.1 GI:14096949

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 2268)

AUTHORS Gearing, D.P. and Busfield, S.J.

TITLE Don-1 gene and polypeptides and uses therefor

JOURNAL Patent: US 6133423-A 31 17-OCT-2000;

FEATURES Location/Qualifiers

source 1. .2268
/organism="unknown"

BASE COUNT 502 a 734 c 701 g 331 t

ORIGIN

Query Match 49.5%; Score 492; DB 6; Length 2268;

Best Local Similarity 92.8%; Pred. No. 1.7e-87;

Matches 516; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Qy 363 CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG 422
| | . | | | | | | | | | | | | | | | | | | | | |

Db 98 CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCGATCCCGGGGAGAA 157

Qy 423 CGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 482
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 158 AGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 217

Qy 483 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 542
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 218 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 277

Qy 543 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 602
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 278 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 337

Qy 603 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 662
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 338 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 397

Qy 663 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 398 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 457

Qy 723 CACCTGTGCATCCTGGTTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 782
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 458 CACCTGTGCATCCTGGTTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 517

Qy 783 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC 842
 |||
 Db 518 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC 577
 Qy 843 AAATGGATTCTTCGGACAGAGATGTTTGGAGAAACTGCCTTTGCGATTGTACATGCCAGA 902
 |||
 Db 578 AAATGGATTCTTCGGACAGAGATGTTTGGAGAAACTGCCTTTGCGATTGTACATGCCAGA 637
 Qy 903 TCCTAAGCAAAGTGTC 918
 |||
 Db 638 TCCTAAGCAAAAAGCC 653

RESULT 15

AB001576

LOCUS AB001576 2188 bp mRNA linear ROD 13-FEB-1999

DEFINITION Rattus sp. mRNA for NTAK alpha2-1p, partial cds.

ACCESSION AB001576

VERSION AB001576.1 GI:2605478

KEYWORDS neural- and thymus-derived activator for ErbB kinases.

SOURCE Rattus sp.

ORGANISM Rattus sp.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (sites)

AUTHORS Higashiyama,S., Horikawa,M., Yamada,K., Ichino,N., Nakano,N.,
 Nakagawa,T., Miyagawa,J., Matsushita,N., Nagatsu,T., Taniguchi,N.
 and Ishiguro,H.

TITLE A novel brain-derived member of the epidermal growth factor family
 that interacts with ErbB3 and ErbB4

JOURNAL J. Biochem. 122 (3), 675-680 (1997)

MEDLINE 98006324

PUBMED 9348101

REFERENCE 2 (bases 1 to 2188)

AUTHORS Ishiguro,H.

TITLE Direct Submission

JOURNAL Submitted (04-MAR-1997) Hiroshi Ishiguro, Fujita Health University,
 ICMS; 1-98, kutsukake-cho, Toyoake, Aichi 470-11, Japan
 (E-mail:hishi@fujita-hu.ac.jp, Tel:0562-93-9393, Fax:0562-93-8831)

FEATURES Location/Qualifiers.

source 1. .2188

/organism="Rattus sp."

/mol_type="mRNA"

/db_xref="taxon:10118"

/cell_line="PC12"

/cell_type="pheochromocytoma"

CDS <1. .1926

/note="neural- and thymus-derived activator for ErbB
 kinases (NTAK); a member of the epidermal growth factor
 (EGF) family"

/codon_start=1

/product="NTAK alpha2-1p"

/protein_id="BAA23348.1"

/db_xref="GI:2605479"

/translation="FFFFKTAFAVPDPNGKNIKKEVGKILCTDCATRPKLKKMKSQTG"

EVGEKQSLKCEAAAGNPQPSYRWFKD GKELNRSRDIRIKYGNRKNRSRLQFNKVKVED
 AGEYVCEAENILGKDTVGRHLVNSVSTTLSSWSGHARKCNETAKSYCVNGGVCYYIE
 GINQLSCKCPNGFFGQRCLEKLPLRLYMPDPKQKAEELYQKRVLTITGICVALLVVGI
 VCVVAYCKTKKQRRQMHHHLRQNMCPAHQNRSLANGPSHPRLDPEEIQMADYISKNP
 ATDHVIRREAETTFSGSHSCSPSHHCSTATPTSSHRHESHTWSLERSESLTSDSQSGI
 MLSSVGTSKCNSPACVEARARRAAAYSQEERRRAAMPYHDSIDSLRDSPHSERYVSA
 LTTPARLSPVDFHYSLATQVPTFEITSPNSAHAVSLPPAAPISYRLAEQQPLL RHPAP
 PGP GPGGADMQRSYDSYYYPAAGPGPRRGACALGGS LGS LPASPFRIPEDDEYETTQ
 ECAPPPPPRPRTRGASRRTSAGPRRWRRSRLNGLAAQRARAARDLSLSLSSGSGCGSAS
 ASDDDADDADGALAAESTPFLGLRAAHDALRSDSPPLCPAADSRTYYSLDSHSTRASS
 RHSRGPPTRAKQDSGPL"

BASE COUNT 515 a 674 c 643 g 356 t
 ORIGIN

Query Match 49.0%; Score 487.4; DB 10; Length 2188;
 Best Local Similarity 90.3%; Pred. No. 1.4e-86;
 Matches 521; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

Qy	337	TTAGTCTTTAAGACGGCCTTTGCCCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAG	396
Db	4	TTTTTTTTTAAGACAGCCTTTGCCCCGGTCGACCCTAACGGCAAAAACATCAAGAAAGAG	63
Qy	397	GTGGGCAAGATCCTGTGCACTGACTGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGC	456
Db	64	GTGGGCAAGATCCTGTGCACTGACTGCGCAACCCGGCCCAAGCTGAAGAAGATGAAGAGT	123
Qy	457	CAGACGGGACAGGTGGGTGAGAAGCAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCC	516
Db	124	CAGACAGGAGAGGTGGGCGAGAAGCAGTCGCTCAAGTGTGAGGCGGCGGGGAACCCC	183
Qy	517	CAGCCTTCCTACCGTTGGTTCAAGGATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGC	576
Db	184	CAGCCCTCCTATCGATGGTTCAAGGACGGCAAGGAGCTCAACCGGAGTCGTGACATTCGC	243
Qy	577	ATCAAATATGGCAACGGCAGAAAGAACTCACGACTACAGTTCAACAAGGTGAAGGTGGAG	636
Db	244	ATCAAGTATGGCAACGGCAGAAAGAACTCACGGCTACAGTTCAACAAAGTGAAGGTGGAG	303
Qy	637	GACGCTGGGGAGTATGTCTGCGAGGCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGC	696
Db	304	GACGCTGGAGAGTACGTCTGTGAGGCTGAGAACATCCTTGGGAAGGACACTGTGAGGGGC	363
Qy	697	CGGCTTTACGTCAACAGCGTGAGCACCACCCTGTCTATCCTGGTCGGGGCAGCCCCGAAG	756
Db	364	CGGCTCCATGTCAACAGTGTGAGCACCCTGTCTCGTCTGGTCGGGGCAGCCCCGAAG	423
Qy	757	TGCAACGAGACAGCCAAGTCCTATTGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGC	816
Db	424	TGCAATGAGACAGCCAAGTCCTACTGTGTGAATGGAGGCGTGTGCTACTACATCGAAGGC	483
Qy	817	ATCAACCAGCTCTCCTGCAAATGTCCAAATGGATTCTTCGGACAGAGATGTTTGGAGAAA	876
Db	484	ATCAACCAACTCTCCTGCAAATGTCCAAACGGATTCTTCGGACAGAGATGTTTGGAGAAA	543
Qy	877	CTGCCTTTGCGATTGTACATGCCAGATCCTAAGCAAA	913
Db	544	CTGCCTTTGCGATTGTACATGCCAGATCCTAAGCAAA	580

Search completed: January 14, 2004, 10:22:56
Job time : 3984.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2004, 07:12:21 ; Search time 321.696 Seconds
(without alignments)
8340.911 Million cell updates/sec

Title: US-09-864-675-1
Perfect score: 994
Sequence: 1 atgaggcgcgcgacccgcgcccc.....caccttggatttgaattaaa 994

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*

- 1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
- 2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
- 3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
- 4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
- 5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
- 6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
- 7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
- 8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
- 9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
- 10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
- 11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
- 12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
- 13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
- 14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
- 15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
- 16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
- 17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
- 18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
- 19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
- 20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
- 21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
- 22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
- 23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
- 24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
- 25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query		DB	ID	Description
	Score	Match Length			
1	994	100.0	994	24	AAS18019
2	900	90.5	1884	19	AAV17814
3	882	88.7	1803	18	AAT87923
4	849	85.4	897	24	AAS18020
5	799.4	80.4	3076	19	AAV43674
6	788	79.3	1863	25	ABS56035
7	738.6	74.3	3441	18	AAT87922
8	547.2	55.1	1607	19	AAV17813
9	535.2	53.8	1561	25	ABS56034
10	492	49.5	2268	19	AAV17816
11	491	49.4	1474	25	ABS56036
12	491	49.4	2266	25	ABS56045
13	490.4	49.3	1476	19	AAV17815
14	464.6	46.7	2467	19	AAV17812
15	455.8	45.9	2442	25	ABS56033
16	424	42.7	1054	24	ABL40993
17	380	38.2	667	18	AAT87924
18	173	17.4	419	24	ABL40994
19	137.2	13.8	1039	23	AAS71393
20	124.6	12.5	480	24	ABL40995
21	122.4	12.3	350	24	ABL40996
22	95.4	9.6	1140	15	AAQ58321
23	95.4	9.6	1140	15	AAQ62840
24	95.4	9.6	1140	16	AAQ74912
25	95.4	9.6	1140	17	AAT48088
26	95.4	9.6	1140	17	AAT31001
27	95.4	9.6	1140	17	AAT06731
28	95.4	9.6	1140	20	AAX81201
29	91.8	9.2	1193	13	AAQ30670
30	91.8	9.2	1193	15	AAQ58303
31	91.8	9.2	1193	15	AAQ62849
32	91.8	9.2	1193	16	AAQ74885
33	91.8	9.2	1193	17	AAT48079
34	91.8	9.2	1193	17	AAT30997
35	91.8	9.2	1193	17	AAT06703
36	87	8.8	300	24	ABL40997
37	79	7.9	1986	20	AAZ32061
38	78.2	7.9	1027	22	AAF80062
39	78.2	7.9	3086	22	AAF80059
40	77.4	7.8	1986	22	AAC90318
41	77.4	7.8	2003	17	AAT48090
42	77.4	7.8	2003	17	AAT30995
43	77.4	7.8	2003	17	AAT06739
44	77.4	7.8	2003	20	AAZ32062
45	77.4	7.8	2003	22	AAC90319

ALIGNMENTS

RESULT 1

AAS18019

ID AAS18019 standard; cDNA; 994 BP.

XX

AC AAS18019;

XX

DT 12-MAR-2002 (first entry)

XX

DE Human cDNA encoding Neuregulin-2alpha, NRG-2alpha.

XX

KW Human; ss; neuregulin-2; NRG-2alpha; NRG-2beta; mitogenesis;

KW cell survival; cell growth; cell differentiation; erbB receptor;

KW cardiomyopathy; ischaemic damage; cardiac trauma; heart failure;

KW atherosclerosis; vascular lesion; vascular hypertension;

KW degenerative congenital vascular disease; myasthenia gravis;

KW neurodegenerative disorder; peripheral neuropathy;

KW sensory nerve fiber neuropathy; motor fiber neuropathy;

KW sensory nerve fiber neuropathy; multiple sclerosis;

KW amyotrophic lateral sclerosis; spinal muscular atrophy; nerve injury;

KW Alzheimer's disease; Parkinson's disease; cerebellar ataxia;

KW spinal cord injury; tumour; neurofibromatosis; transgenic animal.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 1..993

FT /*tag= a

FT /product= "NRG-2alpha"

XX

PN WO200189568-A1.

XX

PD 29-NOV-2001.

XX

PF 23-MAY-2001; 2001WO-US16896.

XX

PR 23-MAY-2000; 2000US-206495P.

XX

PA (CENE-) CENES PHARM INC.

XX

PI Marchionni MA;

XX

DR WPI; 2002-097612/13.

DR P-PSDB; AAU11635.

XX

PT Neuregulin-2 polypeptide and polynucleotide useful for treating

PT multiple sclerosis, spinal muscular atrophy, nerve injury, Alzheimer's

PT disease, by increasing mitogenesis, survival, growth or differentiation

PT of a cell -

XX

PS Claim 57; Fig 6; 79pp; English.

XX

CC The invention relates to a substantially pure neuregulin (NRG)-2

CC polypeptide comprising or consisting of a sequence for human

CC NRG-2alpha or NRG-2beta (clone 2b7) and the polynucleotides encoding

CC the. Also included are a vector expressing the protein, a host cell

CC comprising the vector, a transgenic non-human animal transformed with

CC the vector or having a knockout mutation in one or both NRG-2
 CC alleles and an anti-NRG-2 antibody. Analysis of mutations in NRG-2 in an
 CC individual is useful for diagnosing an increased likelihood of
 CC developing a NRG-2-related disease or condition in a test subject.
 CC NRG-2 is useful for increasing the mitogenesis, survival, growth or
 CC differentiation of a cell (e.g. a neuronal cell), where the cell
 CC expresses an erbB receptor. NRG-2 is useful for treating diseases
 CC and disorders such as cardiomyopathy (preferably degenerative congenital
 CC disease), ischaemic damage, cardiac trauma or heart failure or which
 CC has a condition affecting smooth muscle which include atherosclerosis,
 CC vascular lesion, vascular hypertension, and degenerative congenital
 CC vascular disease, myasthenia gravis, a neurodegenerative disorder,
 CC peripheral neuropathy, a sensory nerve fiber neuropathy, a motor fiber
 CC and a sensory nerve fiber neuropathy, multiple sclerosis, amyotrophic
 CC lateral sclerosis, spinal muscular atrophy, nerve injury, Alzheimer's
 CC disease, Parkinson's disease, cerebellar ataxia, and spinal cord injury.
 CC The antibody is useful for treatment of a tumour comprising inhibiting
 CC proliferation of a tumour cell preferably a glial tumour cell, for
 CC treating of neurofibromatosis by inhibiting glial cell mitogenesis.
 CC The present sequence encodes NRG-2alpha.

XX

SQ Sequence 994 BP; 230 A; 279 C; 304 G; 181 T; 0 other;

Query Match 100.0%; Score 994; DB 24; Length 994;

Best Local Similarity 100.0%; Pred. No. 1.4e-231;

Matches 994; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCGCTCGCCTGC	60
Db	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCGCTCGCCTGC	60
Qy	61	TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Db	61	TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Qy	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Db	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Qy	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Db	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Qy	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Qy	301	CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	301	CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Qy	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Qy	421	TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480

Db	421	TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Qy	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	540
Db	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	540
Qy	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG	600
Db	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG	600
Qy	601	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Db	601	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Qy	661	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Db	661	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Qy	721	ACCACCCTGTCATCCTGGTTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	780
Db	721	ACCACCCTGTCATCCTGGTTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	780
Qy	781	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT	840
Db	781	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT	840
Qy	841	CCAAATGGATTCTTCGGACAGAGATGTTTGGAGAAACTGCCTTTGCGATTGTACATGCCA	900
Db	841	CCAAATGGATTCTTCGGACAGAGATGTTTGGAGAAACTGCCTTTGCGATTGTACATGCCA	900
Qy	901	GATCCTAAGCAAAGTGTCTGTGGGATACACCGGGACAGGTGTCAGCAGTTCGCAATGG	960
Db	901	GATCCTAAGCAAAGTGTCTGTGGGATACACCGGGACAGGTGTCAGCAGTTCGCAATGG	960
Qy	961	TCAACTTCTCCAAGCACCTTGGATTTGAATTAAA	994
Db	961	TCAACTTCTCCAAGCACCTTGGATTTGAATTAAA	994

RESULT 2

AAV17814

ID AAV17814 standard; cDNA; 1884 BP.

XX

AC AAV17814;

XX

DT 17-AUG-1998 (first entry)

XX

DE Homo sapiens don-1 gene splice variant.

XX

KW Murine; don-1 gene; melanoma; treatment; adenocarcinoma;

KW epithelial cell; proliferation; stimulation; treatment; tumours;

KW skin; oesophagus; lung; breast; liver; pancreas; colon; prostate;

KW gastrointestinal tract; uterus; wound healing; transmembrane; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 664..1884
FT /*tag= a
FT /note= "don-1 polypeptide"

XX
PN WO9807736-A1.
XX
PD 26-FEB-1998.
XX
PF 18-AUG-1997; 97WO-US14585.
XX
PR 19-NOV-1996; 96US-0753007.
PR 19-AUG-1996; 96US-0699591.
XX
PA (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.
XX
PI Busfield SJ, Gearing DP;
XX
DR WPI; 1998-169084/15.
DR P-PSDB; AAW48381.
XX

PT Mouse and human don-1 polypeptide(s) - useful for treatment of
PT melanomas and adenocarcinoma(s), and for wound healing
XX

PS Claim 4; Fig 3; 121pp; English.
XX

CC The sequence is that of a human don-1 gene splice variant.
CC Don-1 polypeptides stimulate proliferation of epithelial cells
CC and thus are implicated in melanomas and adenocarcinomas in which
CC epithelial cells proliferate out of control. Compounds that
CC interfere with don-1 mediated cell proliferation can be used
CC in the treatment of tumours such as melanomas and adenocarcinomas
CC of the skin, oesophagus, lung, breast, liver, pancreas,
CC gastrointestinal tract, colon, prostate or uterus. Alternatively,
CC don-1 polypeptides can be used to stimulate epithelial cell
CC proliferation, e.g. for wound healing.
XX

SQ Sequence 1884 BP; 426 A; 607 C; 560 G; 291 T; 0 other;

Query Match 90.5%; Score 900; DB 19; Length 1884;
Best Local Similarity 99.3%; Pred. No. 1.1e-208;
Matches 914; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC 60
|
Db 218 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC 277

Qy 61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
|
Db 278 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 337

Qy 121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
|
Db 338 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 397

Qy 181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
|
Db 398 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 457

Qy	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	458	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	517
Qy	301	CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	518	CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	577
Qy	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	578	CCCCT-GATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGGC	636
Qy	421	TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Db	637	TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	696
Qy	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	540
Db	697	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	756
Qy	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG	600
Db	757	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG	816
Qy	601	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Db	817	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	876
Qy	661	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Db	877	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	936
Qy	721	ACCACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	780
Db	937	ACCACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	996
Qy	781	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT	840
Db	997	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT	1056
Qy	841	CCAAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCA	900
Db	1057	CCAAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCA	1116
Qy	901	GATCCTAAGCAAAGTGTCTCT	920
Db	1117	GATCCTAAGCAAAAGCACCT	1136

RESULT 3

AAT87923

ID AAT87923 standard; cDNA; 1803 BP.

XX

AC AAT87923;

XX

DT 18-DEC-1997 (first entry)

XX
 DE Rat cerebellum derived growth factor 2 cDNA.
 XX
 KW Rat; cerebellum derived growth factor; CDGF2; screening; binding;
 KW modulation; erbB type receptor; identification; indication; risk;
 KW proliferation; differentiation; induction; neuron; hyperplasia;
 KW stem cell culture; intracerebral graft; alleviation; repair;
 KW behavioural defect; nervous system; central; peripheral; nerve;
 KW prosthesis; damage; entubulation; cell survival; treatment;
 KW injury; trauma; ischaemia; ischemia; stroke; infection; disorder;
 KW inflammation; neurodegeneration; disease; Parkinson's;
 KW Huntington's; amyotrophic lateral sclerosis; sensory; retina;
 KW spinocerebellar degeneration; multiple sclerosis; neoplasia;
 KW amalignant glioma; medulloblastoma; neuroectodermal tumour; ds.
 XX
 OS Rattus rattus.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..993
 FT /*tag= a
 FT sig_peptide 1..69
 FT /*tag= b
 FT mat_peptide 70..990
 FT /*tag= c
 FT /product= cerebellum_derived_growth_factor
 XX
 PN WO9709425-A1.
 XX
 PD 13-MAR-1997.
 XX
 PF 09-SEP-1996; 96WO-US14484.
 XX
 PR 08-SEP-1995; 95US-0525864.
 XX
 PA (HARD) HARVARD COLLEGE.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PA (STRD) UNIV LELAND S STANFORD.
 XX
 PI Chang H;
 XX
 DR WPI; 1997-192900/17.
 DR P-PSDB; AAW27537.
 XX
 PT Rat and human cerebellum-derived growth factors - used in the
 PT treatment of neuronal injury and proliferative disorders
 XX
 PS Claim 17; Pages 70-74; 94pp; English.
 XX
 CC The present sequence encodes rat cerebellum derived growth factor 2
 CC (CDGF2), which can be used to screen for modulators of CDGF
 CC binding to erbB type receptors. Identification of a modification or
 CC mutation in a CDGF gene, or aberrant expression of a CDGF gene or
 CC levels of soluble CDGF may be used to indicate the risk of unwanted
 CC cell proliferation or differentiation.
 CC CDGF may be used to induce neuronal differentiation in stem cell
 CC culture, and maintain the integrity of a terminally differentiated
 CC neuronal cell culture, e.g. useful for intracerebral grafting to

CC alleviate behavioural defects. CDGF may also be used in nerve
CC protheses to repair central and peripheral nerve damage, especially
CC where a crushed or severed axon is entubulated by a prosthetic.
CC CDGF may also be used to enhance neuronal cell survival in the
CC central or peripheral nervous system, to treat neurological
CC conditions associated with nervous system injury, e.g. traumatic,
CC chemical or vasal injury and deficits such as ischaemia resulting
CC from stroke, infectious/inflammatory and tumour induced injury,
CC chronic neurodegenerative disease including Parkinson's and
CC Huntingdon's, amyotrophic lateral sclerosis, spinocerebellar
CC degeneration, chronic immunological disease of the nervous system
CC including multiple sclerosis, disorders of the sensory neurons and
CC degenerative diseases of the retina. CDGF may also be used to treat
CC neoplastic or hyperplastic transformations, particularly of the
CC central nervous system, e.g. amalignant gliomas, medulloblastomas
CC and neuroectodermal tumours.

XX

SQ Sequence 1803 BP; 408 A; 549 C; 537 G; 309 T; 0 other;

Query Match 88.7%; Score 882; DB 18; Length 1803;

Best Local Similarity 93.0%; Pred. No. 2.6e-204;

Matches 924; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

```
Qy      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC 60
          |||
Db      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCGATGCTGCTCTTCGGTGTGTCACTCGCCTGC 60

Qy     61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
          |||
Db     61 TACTCGCCCAGCCTCAAGTCCGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120

Qy    121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
          |||
Db    121 GGCAAGGTACAGGGACTGGCCCCGGCAGGCGGTTCCAGCTCTAACAGCACCCGAGAGCCT 180

Qy    181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
          |||
Db    181 CCCGCCTCGGGTCGGGTGGCGCTGGTGAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240

Qy    241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCCAG 300
          |||
Db    241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGCGCGCCGCTCGAAAGGAACCCAG 300

Qy    301 CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
          |||
Db    301 CGCTACATCTTTTTCCTGGAGCCACCGAGCAGCCCTTAGTCTTTAAGACAGCCTTTGCC 360

Qy    361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
          ||
Db    361 CCGGTCGACCCTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420

Qy    421 TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480
          |||
Db    421 TGCGCAACCCGGCCCAAGCTGAAGAAGATGAAGAGTCAGACAGGAGAGGTGGGCGAGAAG 480

Qy    481 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCTACCGTTGGTTCAAG 540
          ||
```


OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..897
FT /*tag= a
FT /product= "NRG-2beta"
XX
PN WO200189568-A1.
XX
PD 29-NOV-2001.
XX
PF 23-MAY-2001; 2001WO-US16896.
XX
PR 23-MAY-2000; 2000US-206495P.
XX
PA (CENE-) CENES PHARM INC.
XX
PI Marchionni MA;
XX
DR WPI; 2002-097612/13.
DR P-PSDB; AAU11636.
XX
PT Neuregulin-2 polypeptide and polynucleotide useful for treating
PT multiple sclerosis, spinal muscular atrophy, nerve injury, Alzheimer's
PT disease, by increasing mitogenesis, survival, growth or differentiation
PT of a cell -
XX
PS Claim 57; Fig 8; 79pp; English.
XX
CC The invention relates to a substantially pure neuregulin (NRG)-2
CC polypeptide comprising or consisting of a sequence for human
CC NRG-2alpha or NRG-2beta (clone 2b7) and the polynucleotides encoding
CC the. Also included are a vector expressing the protein, a host cell
CC comprising the vector, a transgenic non-human animal transformed with
CC the vector or having a knockout mutation in one or both NRG-2
CC alleles and an anti-NRG-2 antibody. Analysis of mutations in NRG-2 in an
CC individual is useful for diagnosing an increased likelihood of
CC developing a NRG-2-related disease or condition in a test subject.
CC NRG-2 is useful for increasing the mitogenesis, survival, growth or
CC differentiation of a cell (e.g. a neuronal cell), where the cell
CC expresses an erbB receptor. NRG-2 is useful for treating diseases
CC and disorders such as cardiomyopathy (preferably degenerative congenital
CC disease), ischaemic damage, cardiac trauma or heart failure or which
CC has a condition affecting smooth muscle which include atherosclerosis,
CC vascular lesion, vascular hypertension, and degenerative congenital
CC vascular disease, myasthenia gravis, a neurodegenerative disorder,
CC peripheral neuropathy, a sensory nerve fiber neuropathy, a motor fiber
CC and a sensory nerve fiber neuropathy, multiple sclerosis, amyotrophic
CC lateral sclerosis, spinal muscular atrophy, nerve injury, Alzheimer's
CC disease, Parkinson's disease, cerebellar ataxia, and spinal cord injury.
CC The antibody is useful for treatment of a tumour comprising inhibiting
CC proliferation of a tumour cell preferably a glial tumour cell, for
CC treating of neurofibromatosis by inhibiting glial cell mitogenesis.
CC The present sequence encodes NRG-2beta.
XX
SQ Sequence 897 BP; 200 A; 261 C; 282 G; 154 T; 0 other;

Query Match 85.4%; Score 849; DB 24; Length 897;
 Best Local Similarity 98.3%; Pred. No. 2.1e-196;
 Matches 858; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC	60
Db	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC	60
Qy	61	TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Db	61	TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Qy	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Db	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Qy	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGCCGCTCCGGAGCGGG	240
Db	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGCCGCTCCGGAGCGGG	240
Qy	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Qy	301	CGCTACATCTTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	301	CGCTACATCTTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Qy	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Qy	421	TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Db	421	TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Qy	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCAGCCTTCCTACCGTTGGTTCAAG	540
Db	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCAGCCTTCCTACCGTTGGTTCAAG	540
Qy	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTGCGATCAAATATGGCAACGGCAGAAAG	600
Db	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTGCGATCAAATATGGCAACGGCAGAAAG	600
Qy	601	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Db	601	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Qy	661	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Db	661	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Qy	721	ACCACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	780
Db	721	ACCACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	780
Qy	781	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT	840

```

Db      781 TGGCTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT 840
Qy      841 CCAAATGGATTCTTCGGACAGAGATGTTTGGAG 873
      ||      |||| | ||| | || ||| | ||
Db      841 CCTGTGGGATACACCGGGGACAGGTGTCAGCAG 873

```

RESULT 5

AAV43674

ID AAV43674 standard; cDNA; 3076 BP.

XX

AC AAV43674;

XX

DT 29-SEP-1998 (first entry)

XX

DE Receptor type tyrosine kinase ErbB ligand encoding cDNA.

XX

KW Receptor type tyrosine kinase ErbB; ligand; diagnostic agent;
KW nervous disease; cancer; ss.

XX

OS Rattus sp.

XX

FH Key Location/Qualifiers

FT CDS 232..2814

FT /*tag= a

FT /product= "ligand of receptor type tyrosine kinase ErbB"

XX

PN JP10179166-A.

XX

PD 07-JUL-1998.

XX

PF 25-DEC-1996; 96JP-0356998.

XX

PR 25-DEC-1996; 96JP-0356998.

XX

PA (HIGA/) HIGASHIYAMA S.

XX

DR WPI; 1998-430952/37.

DR P-PSDB; AAW63700.

XX

PT Gene coding the ligand of the tyrosine kinase ErbB receptor - useful
PT for diagnosing and treating nervous diseases and cancer

XX

PS Examples; Pages 9-13; 17pp; Japanese.

XX

CC This cDNA encodes the ligand of receptor type tyrosine kinase ErbB. A
CC prokaryotic or eukaryotic host cell transformed by a recombinant vector
CC containing the encoding DNA can be used for the recombinant production of
CC the protein. The invention provides a method for inhibiting the formation
CC of the ligand of receptor type tyrosine kinase ErbB in an animal using
CC an antibody recognizing the protein. The ligand of the tyrosine kinase
CC ErbB receptor and associated materials can be used for treating or
CC diagnosing nervous diseases and cancers.

XX

SQ Sequence 3076 BP; 673 A; 996 C; 944 G; 463 T; 0 other;

Query Match 80.4%; Score 799.4; DB 19; Length 3076;
Best Local Similarity 92.2%; Pred. No. 3.4e-184;
Matches 842; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

```
Qy      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC 60
          |||
Db      556 ATGAGGCGCGACCCGGCCCCCGGCTCCTCGATGCTGCTCTTCGGTGTGTCACTCGCCTGC 615

Qy      61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
          |||
Db      616 TACTCGCCCAGCCTCAAGTCCGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 675

Qy     121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
          |||
Db     676 GGCAAGGTACAGGGACTGGCCCCGGCAGGCGGTTCCAGCTCTAACAGCACCCGAGAGCCT 735

Qy     181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
          |||
Db     736 CCCGCCTCGGGTCGGGTGGCGCTGGTGAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 795

Qy     241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
          |||
Db     796 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGCGCGCCGCTCGAAAGGAACCAG 855

Qy     301 CGCTACATCTTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
          |||
Db     856 CGCTACATCTTTTTCTGGAGCCACCGAGCAGCCCTTAGTTTTTAAGACAGCCTTTGCC 915

Qy     361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
          ||
Db     916 CCGGTCGACCCTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 975

Qy     421 TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480
          |||
Db     976 TGCGCAACCCGGCCCCAAGCTGAAGAAGATGAAGAGTCAGACAGGAGAGGTGGGCGAGAAG 1035

Qy     481 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 540
          ||
Db    1036 CAGTCGCTCAAGTGTGAGGCGGCGGCGGGGAACCCCGAGCCCTCCTATCGATGGTTCAAG 1095

Qy     541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTGCGATCAAATATGGCAACGGCAGAAAG 600
          ||
Db    1096 GACGGCAAGGAGCTCAACCGGAGTCGTGACATTGCGATCAAGTATGGCAACGGCAGAAAG 1155

Qy     601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660
          |||
Db    1156 AACTCACGGCTACAGTTCAACAAAGTGAAGGTGGAGGACGCTGGAGAGTACGTCTGTGAG 1215

Qy     661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720
          ||
Db    1216 GCTGAGAACATCCTTGGGAAGGACACTGTGAGGGGCGGGCTCCATGTCAACAGTGTGAGC 1275

Qy     721 ACCACCCTGTCATCCTGGTCGGGGCACGCCCCGAAGTGCAACGAGACAGCCAAGTCCTAT 780
          |||
Db    1276 ACCACTCTGTCTGCTGGTCGGGGCACGCCCCGAAGTGCAATGAGACAGCCAAGTCCTAC 1335

Qy     781 TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT 840
```

```

      || || ||||| ||||| ||||| ||||| |||||
Db      1336 TGTGTGAATGGAGGCGTGTGCTACTACATCGAAGGCATCAACCAACTCTCCTGCAAATGT 1395

Qy      841 CCAAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCA 900
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1396 CCAAACGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCA 1455

Qy      901 GATCCTAAGCAAA 913
      ||||| |||||
Db      1456 GATCCTAAGCAAA 1468

```

RESULT 6

ABS56035

ID ABS56035 standard; cDNA; 1863 BP.

XX

AC ABS56035;

XX

DT 14-JAN-2003 (first entry)

XX

DE cDNA encoding human membrane-bound splice variant of Don-1.

XX

KW Human; Don-1; epidermal growth factor; EGF; neuregulin;

KW glycoprotein ligand; cell proliferation; cell proliferative disorder;

KW carcinoma; adenocarcinoma cell; myeloma; cell differentiation;

KW cell survival; epithelial cell; wound healing; tumour formation;

KW brain; vulnerary; cytostatic; gene therapy; gene; ss.

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT	CDS	643..1863
----	-----	-----------

FT		/*tag= a
----	--	----------

FT		/partial
----	--	----------

FT		/product= "Membrane-bound splice variant of Don-1"
----	--	--

FT		/note= "This sequence lacks a stop codon"
----	--	---

XX

PN US2002127594-A1.

XX

PD 12-SEP-2002.

XX

PF 12-MAR-2002; 2002US-0096241.

XX

PR 22-JUN-2000; 2000US-0599789.

XX

PA (GEAR/) GEARING D P.

PA (BUSF/) BUSFIELD S J.

XX

PI Gearing DP, Busfield SJ;

XX

DR WPI; 2003-039584/03.

DR P-PSDB; ABG71638.

XX

PT Novel Don-1 polypeptide useful for stimulating proliferation of cells,

PT for identifying proteins that interact with Don-1, and for regulating

PT tumour formation and progression in brain -

XX

Db	618	TGCGCCACCCGGCCCAAGTTGA--AAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	675
Qy	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	540
Db	676	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	735
Qy	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG	600
Db	736	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG	795
Qy	601	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Db	796	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	855
Qy	661	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Db	856	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	915
Qy	721	ACCACCCTGTCATCCTGGTCGGGGCAGCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	780
Db	916	ACCACCCTGTCATCCTGGTCGGGGCAGCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	975
Qy	781	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT	840
Db	976	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT	1035
Qy	841	CCAAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCA	900
Db	1036	CCAAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCA	1095
Qy	901	GATCCTAAGCAAAGTGTCTCT	920
Db	1096	GATCCTAAGCAAAGCACCT	1115

RESULT 7

AAT87922

ID AAT87922 standard; cDNA; 3441 BP.

XX

AC AAT87922;

XX

DT 18-DEC-1997 (first entry)

XX

DE Rat cerebellum derived growth factor 1 cDNA.

XX

KW Rat; cerebellum derived growth factor; CDGF1; screening; binding;

KW modulation; erbB type receptor; identification; indication; risk;

KW proliferation; differentiation; induction; neuron; hyperplasia;

KW stem cell culture; intracerebral graft; alleviation; repair;

KW behavioural defect; nervous system; central; peripheral; nerve;

KW prosthesis; damage; entubulation; cell survival; treatment;

KW injury; trauma; ischaemia; ischemia; stroke; infection; disorder;

KW inflammation; neurodegeneration; disease; Parkinson's;

KW Huntington's; amyotrophic lateral sclerosis; sensory; retina;

KW spinocerebellar degeneration; multiple sclerosis; neoplasia;

KW amalignant glioma; medulloblastoma; neuroectodermal tumour; ds.

XX

OS Rattus rattus.
 XX
 FH Key Location/Qualifiers
 FT CDS 180..2444
 FT /*tag= a
 FT sig_peptide 180..248
 FT /*tag= b
 FT mat_peptide 249..2441
 FT /*tag= c
 FT /product= cerebellum_derived_growth_factor
 XX
 PN WO9709425-A1.
 XX
 PD 13-MAR-1997.
 XX
 PF 09-SEP-1996; 96WO-US14484.
 XX
 PR 08-SEP-1995; 95US-0525864.
 XX
 PA (HARD) HARVARD COLLEGE.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PA (STRD) UNIV LELAND S STANFORD.
 XX
 PI Chang H;
 XX
 DR WPI; 1997-192900/17.
 DR P-PSDB; AAW27536.
 XX
 PT Rat and human cerebellum-derived growth factors - used in the
 PT treatment of neuronal injury and proliferative disorders
 XX
 PS Claim 17; Pages 63-66; 94pp; English.
 XX
 CC The present sequence encodes rat cerebellum derived growth factor 1
 CC (CDGF1), which can be used to screen for modulators of CDGF
 CC binding to erbB type receptors. Identification of a modification or
 CC mutation in a CDGF gene, or aberrant expression of a CDGF gene or
 CC levels of soluble CDGF may be used to indicate the risk of unwanted
 CC cell proliferation or differentiation.
 CC CDGF may be used to induce neuronal differentiation in stem cell
 CC culture, and maintain the integrity of a terminally differentiated
 CC neuronal cell culture, e.g. useful for intracerebral grafting to
 CC alleviate behavioural defects. CDGF may also be used in nerve
 CC prostheses to repair central and peripheral nerve damage, especially
 CC where a crushed or severed axon is entubulated by a prosthetic.
 CC CDGF may also be used to enhance neuronal cell survival in the
 CC central or peripheral nervous system, to treat neurological
 CC conditions associated with nervous system injury, e.g. traumatic,
 CC chemical or vasal injury and deficits such as ischaemia resulting
 CC from stroke, infectious/inflammatory and tumour induced injury,
 CC chronic neurodegenerative disease including Parkinson's and
 CC Huntingdon's, amyotrophic lateral sclerosis, spinocerebellar
 CC degeneration, chronic immunological disease of the nervous system
 CC including multiple sclerosis, disorders of the sensory neurons and
 CC degenerative diseases of the retina. CDGF may also be used to treat
 CC neoplastic or hyperplastic transformations, particularly of the
 CC central nervous system, e.g. amalignant gliomas, medulloblastomas

CC and neuroectodermal tumours.

XX

SQ Sequence 3441 BP; 777 A; 1057 C; 1015 G; 592 T; 0 other;

Query Match 74.3%; Score 738.6; DB 18; Length 3441;
Best Local Similarity 90.4%; Pred. No. 2e-169;
Matches 789; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

```
Qy      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCCTGC 60
          |||
Db      180 ATGAGGCGCGACCCGGCCCCCGGCTTCTCGATGCTGCTCTTCGGTGTGTCACTCGCCTGC 239

Qy      61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
          |||
Db      240 TACTCGCCCAGCCTCAAGTCCGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 299

Qy     121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
          |||
Db     300 GGCAAGGTACAGGGACTGGCCCCGGCAGGCGGTTCCAGCTCTAACAGCACCCGAGAGCCT 359

Qy     181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
          |||
Db     360 CCCGCCTCGGGTCGGGTGGCGCTGGTGAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 419

Qy     241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
          |||
Db     420 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGCGCGCCGCTCGAAAGGAACCAG 479

Qy     301 CGCTACATCTTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
          |||
Db     480 CGCTACATCTTTTTCTGGAGCCACCGAGCAGCCCTTAGTTTTTAAGACAGCCTTTGCC 539

Qy     361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
          ||
Db     540 CCGGTCGACCCTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 599

Qy     421 TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480
          |||
Db     600 TGCGCAACCCGGCCCCAAGCTGAAGAAGATGAAGAGTCAGACAGGAGAGGTGGGCGAGAAG 659

Qy     481 CAATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCTACCGTTGGTTCAAG 540
          ||
Db     660 CAGTCGCTCAAGTGTGAGGCGGCGGGGAACCCCGAGCCCTCCTATCGATGGTTCAAG 719

Qy     541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG 600
          ||
Db     720 GACGGCAAGGAGCTCAACCGGAGTCGTGACATTTCGCATCAAGTATGGCAACGGCAGAAAG 779

Qy     601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660
          |||
Db     780 AACTCACGGCTACAGTTCAACAAAGTGAAGGTGGAGGACGCTGGAGAGTACGTCTGTGAG 839

Qy     661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720
          ||
Db     840 GCTGAGAACATCCTTGGGAAGGACACTGTGAGGGGCCGGCTCCATGTCAACAGTGTGAGC 899

Qy     721 ACCACCCTGTCATCCTGGTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 780
```



```

          ||||| ||||| ||||||||||||||||||||||||||||| ||||||||||||||||
Db      900 ACCACTCTGTCGTCCTGGTCGGGGGCACGCCCGGAAGTGCAATGAGACAGCCAAGTCCTAC 959

Qy      781 TGCCTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT 840
          || || ||||||||||| ||||||||||||| ||||||||||| ||||||||||| |||
Db      960 TGTGTGAATGGAGGCGTGTGCTACTACATCGAAGGCATCAACCAACTCTCCTGCAAGTGT 1019

Qy      841 CCAAATGGATTCTTCGGACAGAGATGTTTGGAG 873
          || ||||| | ||| | || ||| | ||
Db      1020 CCTGTGGGATACACCGGGGACAGGTGTCAGCAG 1052

```

RESULT 8

AAV17813

ID AAV17813 standard; cDNA; 1607 BP.

XX

AC AAV17813;

XX

DT 17-AUG-1998 (first entry)

XX

DE Mus musculus don-1 gene splice variant.

XX

KW Murine; don-1 gene; melanoma; treatment; adenocarcinoma;

KW epithelial cell; proliferation; stimulation; treatment; tumours;

KW skin; oesophagus; lung; breast; liver; pancreas; colon; prostate;

KW gastrointestinal tract; uterus; wound healing; secreted protein; ss.

XX

OS Mus musculus.

XX

FH Key Location/Qualifiers

FT CDS 79..624

FT /*tag= a

FT /note= "secreted don-1 polypeptide"

XX

PN WO9807736-A1.

XX

PD 26-FEB-1998.

XX

PF 18-AUG-1997; 97WO-US14585.

XX

PR 19-NOV-1996; 96US-0753007.

PR 19-AUG-1996; 96US-0699591.

XX

PA (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.

XX

PI Busfield SJ, Gearing DP;

XX

DR WPI; 1998-169084/15.

DR P-PSDB; AAW48380.

XX

PT Mouse and human don-1 polypeptide(s) - useful for treatment of

PT melanomas and adenocarcinoma(s), and for wound healing

XX

PS Claim 4; Fig 2; 12lpp; English.

XX

CC The sequence is that of a murine don-1 gene splice variant.

CC Don-1 polypeptides stimulate proliferation of epithelial cells

Db

602 CAAGCACCTTGGATTGAATTGAA 625

RESULT 9

ABS56034

ID ABS56034 standard; cDNA; 1561 BP.

XX

AC ABS56034;

XX

DT 14-JAN-2003 (first entry)

XX

DE cDNA encoding murine secreted splice variant of Don-1.

XX

KW Murine; Don-1; epidermal growth factor; EGF; neuregulin; mouse;
KW glycoprotein ligand; cell proliferation; cell proliferative disorder;
KW carcinoma; adenocarcinoma cell; myeloma; cell differentiation;
KW cell survival; epithelial cell; wound healing; tumour formation;
KW brain; vulnerary; cytostatic; gene therapy; chromosome 18; gene; ss.

XX

OS Mus sp.

XX

FH Key Location/Qualifiers

FT CDS 78..623

FT /*tag= a

FT /product= "Secreted splice variant of Don-1"

XX

PN US2002127594-A1.

XX

PD 12-SEP-2002.

XX

PF 12-MAR-2002; 2002US-0096241.

XX

PR 22-JUN-2000; 2000US-0599789.

XX

PA (GEAR/) GEARING D P.

PA (BUSF/) BUSFIELD S J.

XX

PI Gearing DP, Busfield SJ;

XX

DR WPI; 2003-039584/03.

DR P-PSDB; ABG71637.

XX

PT Novel Don-1 polypeptide useful for stimulating proliferation of cells,
PT for identifying proteins that interact with Don-1, and for regulating
PT tumour formation and progression in brain -

XX

PS Claim 4; Fig 2; 66pp; English.

XX

CC The present invention relates to the isolation of a novel gene
CC called Don-1, and alternate splice variants of Don-1, which are
CC related to epidermal growth factors (EGF) such as neuregulins.
CC Don-1 polypeptides are glycoprotein ligands. Both murine and human
CC Don-1 sequences are cloned. The mouse Don-1 gene maps to chromosome 18.
CC Don-1 polypeptides are useful for stimulating proliferation of a cell.
CC Antibodies to Don-1 polypeptides are useful for detecting Don-1
CC in a sample. The Don-1 polypeptides are useful for treating and
CC diagnosing cell proliferative disorders and play a role in the

Db

|||||
601 CAAGCACCTTGGATTGAATTGAA 624

RESULT 10

AAV17816

ID AAV17816 standard; cDNA; 2268 BP.

XX

AC AAV17816;

XX

DT 17-AUG-1998 (first entry)

XX

DE Homo sapiens don-1 gene splice variant.

XX

KW Murine; don-1 gene; melanoma; treatment; adenocarcinoma;

KW epithelial cell; proliferation; stimulation; treatment; tumours;

KW skin; oesophagus; lung; breast; liver; pancreas; colon; prostate;

KW gastrointestinal tract; uterus; wound healing; transmembrane; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 69..2012

FT /*tag= a

FT /note= "don-1 polypeptide"

XX

PN WO9807736-A1.

XX

PD 26-FEB-1998.

XX

PF 18-AUG-1997; 97WO-US14585.

XX

PR 19-NOV-1996; 96US-0753007.

PR 19-AUG-1996; 96US-0699591.

XX

PA (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.

XX

PI Busfield SJ, Gearing DP;

XX

DR WPI; 1998-169084/15.

DR P-PSDB; AAW48383.

XX

PT Mouse and human don-1 polypeptide(s) - useful for treatment of

PT melanomas and adenocarcinoma(s), and for wound healing

XX

PS Claim 4; Fig 7; 121pp; English.

XX

CC The sequence is that of a human don-1 gene splice variant.

CC Don-1 polypeptides stimulate proliferation of epithelial cells

CC and thus are implicated in melanomas and adenocarcinomas in which

CC epithelial cells proliferate out of control. Compounds that

CC interfere with don-1 mediated cell proliferation can be used

CC in the treatment of tumours such as melanomas and adenocarcinomas

CC of the skin, oesophagus, lung, breast, liver, pancreas,

CC gastrointestinal tract, colon, prostate or uterus. Alternatively,

CC don-1 polypeptides can be used to stimulate epithelial cell

CC proliferation, e.g. for wound healing.

Sequence 2268 BP; 502 A; 735 C; 700 G; 331 T; 0 other;

Query Match 49.5%; Score 492; DB 19; Length 2268;
Best Local Similarity 92.8%; Pred. No. 1.3e-109;
Matches 516; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Qy	363	CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG	422
Db	98	CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCGATCCCGGGGAGAA	157
Qy	423	CGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA	482
Db	158	AGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA	217
Qy	483	ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA	542
Db	218	ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA	277
Qy	543	TGGCAAGGAGCTCAACCGCAGCCGAGACATTGCGATCAAATATGGCAACGGCAGAAAGAA	602
Db	278	TGGCAAGGAGCTCAACCGCAGCCGAGACATTGCGATCAAATATGGCAACGGCAGAAAGAA	337
Qy	603	CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC	662
Db	338	CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC	397
Qy	663	CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC	722
Db	398	CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC	457
Qy	723	CACCCTGTCATCCTGGTCGGGGCACGCCCGBAAGTGCAACGAGACAGCCAAGTCCTATTG	782
Db	458	CACCCTGTCATCCTGGTCGGGGCACGCCCGBAAGTGCAACGAGACAGCCAAGTCCTATTG	517
Qy	783	CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC	842
Db	518	CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC	577
Qy	843	AAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCAGA	902
Db	578	AAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCAGA	637
Qy	903	TCCTAAGCAAAGTGTC	918
Db	638	TCCTAAGCAAAAAGCC	653

ABS56036

ID ABS56036 standard; cDNA; 1474 BP.

XX

AC ABS56036;

XX

DT 14-JAN-2003 (first entry)

XX

DE cDNA encoding human second splice variant of Don-1.

XX
 KW Human; Don-1; epidermal growth factor; EGF; neuregulin;
 KW glycoprotein ligand; cell proliferation; cell proliferative disorder;
 KW carcinoma; adenocarcinoma cell; myeloma; cell differentiation;
 KW cell survival; epithelial cell; wound healing; tumour formation;
 KW brain; vulnerary; cytostatic; gene therapy; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 68..1473
 FT /*tag= a
 FT /partial
 FT /product= "Second splice variant of Don-1"
 FT /note= "This sequence lacks a stop codon"
 FT /transl_except= (pos:107..108, aa:Lys)
 FT /note= "This codon has an apparent 1 nucleotide
 FT deletion which alters the reading frame"
 XX
 PN US2002127594-A1.
 XX
 PD 12-SEP-2002.
 XX
 PF 12-MAR-2002; 2002US-0096241.
 XX
 PR 22-JUN-2000; 2000US-0599789.
 XX
 PA (GEAR/) GEARING D P.
 PA (BUSF/) BUSFIELD S J.
 XX
 PI Gearing DP, Busfield SJ;
 XX
 DR WPI; 2003-039584/03.
 DR P-PSDB; ABG71639.
 XX
 PT Novel Don-1 polypeptide useful for stimulating proliferation of cells,
 PT for identifying proteins that interact with Don-1, and for regulating
 PT tumour formation and progression in brain -
 XX
 PS Claim 4; Fig 4; 66pp; English.
 XX
 CC The present invention relates to the isolation of a novel gene
 CC called Don-1, and alternate splice variants of Don-1, which are
 CC related to epidermal growth factors (EGF) such as neuregulins.
 CC Don-1 polypeptides are glycoprotein ligands. Both murine and human
 CC Don-1 sequences are cloned. The mouse Don-1 gene maps to chromosome 18.
 CC Don-1 polypeptides are useful for stimulating proliferation of a cell.
 CC Antibodies to Don-1 polypeptides are useful for detecting Don-1
 CC in a sample. The Don-1 polypeptides are useful for treating and
 CC diagnosing cell proliferative disorders and play a role in the
 CC proliferation of carcinomas e.g. adenocarcinoma, myeloma, in cell
 CC differentiation, proliferation and survival. The polypeptides are
 CC also useful for inhibiting proliferation of adenocarcinoma cells,
 CC for stimulating the proliferation of cells such as epithelial cells
 CC to promote wound healing, for identifying proteins that interact
 CC with Don-1, and for regulating tumour formation and progression in
 CC the brain. The polynucleotide sequences encoding Don-1 may be used

KW glycoprotein ligand; cell proliferation; cell proliferative disorder;
 KW carcinoma; adenocarcinoma cell; myeloma; cell differentiation;
 KW cell survival; epithelial cell; wound healing; tumour formation;
 KW brain; vulnerary; cytostatic; gene therapy; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 68..2010
 FT /*tag= a
 FT /product= "Third splice variant of Don-1"
 FT /transl_except= (pos:107..108, aa:Lys)
 FT /note= "This codon has an apparent 1 nucleotide
 FT deletion which alters the reading frame"
 FT /transl_except= (pos:994..996, aa:Thr)
 XX
 PN US2002127594-A1.
 XX
 PD 12-SEP-2002.
 XX
 PF 12-MAR-2002; 2002US-0096241.
 XX
 PR 22-JUN-2000; 2000US-0599789.
 XX
 PA (GEAR/) GEARING D P.
 PA (BUSF/) BUSFIELD S J.
 XX
 PI Gearing DP, Busfield SJ;
 XX
 DR WPI; 2003-039584/03.
 DR P-PSDB; ABG71644.
 XX
 PT Novel Don-1 polypeptide useful for stimulating proliferation of cells,
 PT for identifying proteins that interact with Don-1, and for regulating
 PT tumour formation and progression in brain -
 XX
 PS Claim 4; Fig 7; 66pp; English.
 XX
 CC The present invention relates to the isolation of a novel gene
 CC called Don-1, and alternate splice variants of Don-1, which are
 CC related to epidermal growth factors (EGF) such as neuregulins.
 CC Don-1 polypeptides are glycoprotein ligands. Both murine and human
 CC Don-1 sequences are cloned. The mouse Don-1 gene maps to chromosome 18.
 CC Don-1 polypeptides are useful for stimulating proliferation of a cell.
 CC Antibodies to Don-1 polypeptides are useful for detecting Don-1
 CC in a sample. The Don-1 polypeptides are useful for treating and
 CC diagnosing cell proliferative disorders and play a role in the
 CC proliferation of carcinomas e.g. adenocarcinoma, myeloma, in cell
 CC differentiation, proliferation and survival. The polypeptides are
 CC also useful for inhibiting proliferation of adenocarcinoma cells,
 CC for stimulating the proliferation of cells such as epithelial cells
 CC to promote wound healing, for identifying proteins that interact
 CC with Don-1, and for regulating tumour formation and progression in
 CC the brain. The polynucleotide sequences encoding Don-1 may be used
 CC in gene therapy. The present sequence encodes human third
 CC splice variant of Don-1.
 XX


```

XX
OS   Homo sapiens.
XX
FH   Key                Location/Qualifiers
FT   CDS                69..1475
FT                               /*tag= a
FT                               /note= "don-1 polypeptide"
XX
PN   WO9807736-A1.
XX
PD   26-FEB-1998.
XX
PF   18-AUG-1997;      97WO-US14585.
XX
PR   19-NOV-1996;      96US-0753007.
PR   19-AUG-1996;      96US-0699591.
XX
PA   (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.
XX
PI   Busfield SJ,  Gearing DP;
XX
DR   WPI; 1998-169084/15.
DR   P-PSDB; AAW48382.
XX
PT   Mouse and human don-1 polypeptide(s) - useful for treatment of
PT   melanomas and adenocarcinoma(s), and for wound healing
XX
PS   Claim 4; Fig 4; 12lpp; English..
XX
CC   The sequence is that of a human don-1 gene splice variant.
CC   Don-1 polypeptides stimulate proliferation of epithelial cells
CC   and thus are implicated in melanomas and adenocarcinomas in which
CC   epithelial cells proliferate out of control. Compounds that
CC   interfere with don-1 mediated cell proliferation can be used
CC   in the treatment of tumours such as melanomas and adenocarcinomas
CC   of the skin, oesophagus, lung, breast, liver, pancreas,
CC   gastrointestinal tract, colon, prostate or uterus. Alternatively,
CC   don-1 polypeptides can be used to stimulate epithelial cell
CC   proliferation, e.g. for wound healing.
XX
SQ   Sequence 1476 BP; 335 A; 475 C; 450 G; 216 T; 0 other;

      Query Match          49.3%;  Score 490.4;  DB 19;  Length 1476;
      Best Local Similarity 92.6%;  Pred. No. 2.9e-109;
      Matches 515;  Conservative 0;  Mismatches 41;  Indels 0;  Gaps 0;

Qy      363 CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG 422
      || || | || | | | | || | | | | | | | | | |
Db      98  CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCGATCCCGGGGAGAA 157

Qy      423 CGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 482
      |||||
Db      158 AGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 217

Qy      483 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 542
      |||||
Db      218 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 277

```

Qy 543 TGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAA 602
 |||
 Db 278 TGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAA 337
 Qy 603 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 662
 |||
 Db 338 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 397
 Qy 663 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722
 |||
 Db 398 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 457
 Qy 723 CACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 782
 |||
 Db 458 CACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 517
 Qy 783 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC 842
 |||
 Db 518 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC 577
 Qy 843 AAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCAGA 902
 |||
 Db 578 AAATGGATTCTTCGCACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCAGA 637
 Qy 903 TCCTAAGCAAAGTGTC 918
 |||
 Db 638 TCCTAAGCAAAAAGCC 653

RESULT 14

AAV17812

ID AAV17812 standard; cDNA; 2467 BP.

XX

AC AAV17812;

XX

DT 17-AUG-1998 (first entry)

XX

DE Mus musculus don-1 gene splice variant.

XX

KW Murine; don-1 gene; melanoma; treatment; adenocarcinoma;

KW epithelial cell; proliferation; stimulation; treatment; tumours;

KW skin; oesophagus; lung; breast; liver; pancreas; colon; prostate;

KW gastrointestinal tract; uterus; wound healing; transmembrane; ss.

XX

OS Mus musculus.

XX

FH Key Location/Qualifiers

FT CDS 79..1896

FT /*tag= a

FT /note= "transmembrane don-1 polypeptide"

XX

PN WO9807736-A1.

XX

PD 26-FEB-1998.

XX

PF 18-AUG-1997; 97WO-US14585.

```
XX PR 19-NOV-1996; 96US-0753007.
PR 19-AUG-1996; 96US-0699591.
XX
PA (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.
XX
PI Busfield SJ, Gearing DP;
XX
DR WPI; 1998-169084/15.
DR P-PSDB; AAW48379.
XX
PT Mouse and human don-1 polypeptide(s) - useful for treatment of
PT melanomas and adenocarcinoma(s), and for wound healing
XX
PS Claim 4; Fig 1; 121pp; English.
XX
CC The sequence is that of a murine don-1 gene splice variant.
CC Don-1 polypeptides stimulate proliferation of epithelial cells
CC and thus are implicated in melanomas and adenocarcinomas in which
CC epithelial cells proliferate out of control. Compounds that
CC interfere with don-1 mediated cell proliferation can be used
CC in the treatment of tumours such as melanomas and adenocarcinomas
CC of the skin, oesophagus, lung, breast, liver, pancreas,
CC gastrointestinal tract, colon, prostate or uterus. Alternatively,
CC don-1 polypeptides can be used to stimulate epithelial cell
CC proliferation, e.g. for wound healing.
XX
SQ Sequence 2467 BP; 592 A; 752 C; 706 G; 417 T; 0 other;

Query Match          46.7%; Score 464.6; DB 19; Length 2467;
Best Local Similarity 91.0%; Pred. No. 6.2e-103;
Matches 494; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY      371 CCAACGGCAA AAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCCC 430
        | |||||
Db       2   CTAACGGCAA AAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCCC 61

QY      431 GGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAATCGCTGA 490
        |||||
Db       62 GGCCCAAGCTGAAGAAGATGAAGAGCCAGACAGGAGAGGTGGGTGAGAAGCAGTCGCTCA 121

QY      491 AGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACC GTTGTTCAAGGATGGCAAGG 550
        |||||
Db      122 AGTGTGAGGCAGCGGCGGGAAACCCCCAGCCCTCCTATCGCTGGTTCAAGGATGGCAAGG 181

QY      551 AGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAAGAACTCACGAC 610
        | |||||
Db      182 AACTCAACCGGAGTCGTGATATTTCGCATCAAGTATGGCAATGTCAGAAAAGAACTCACGGC 241

QY      611 TACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGCCGAGAACA 670
        |||||
Db      242 TACAGTTCAACAAAGTGAGCGTGGAGGATGCCGGGGAGTACGTCTGTGAGGCCGAGAACA 301

QY      671 TCCTGGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCACCACCCTGT 730
        |||||
Db      302 TCCTTGGGAAGGACACCGTGAGGGGCGGACTCCATGTCAACAGCGTGACCACCACCTCTGT 361
```



```

Qy      791 GAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCCAAATGGAT 850
        ||||||| |||||||||||||||||||||||||||||||||||||||| ||||
Db      421 GAGGCGTGTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCCAAACGGAT 480

Qy      851 TCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCAGATCCTAAGC 910
        ||||||| |||||||||||||||||||||||||||||||||||||||| ||||
Db      481 TCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCAGATCCTAAGC 540

Qy      911 AAA 913
        |||
Db      541 AAA 543

```

Search completed: January 14, 2004, 08:16:37
 Job time : 324.696 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2004, 07:16:01 ; Search time 73.5907 Seconds
(without alignments)
5961.825 Million cell updates/sec

Title: US-09-864-675-1
Perfect score: 994
Sequence: 1 atgaggcgcgacccggcccc.....caccttgatttgaattaaa 994

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8						
Result	Query					
No.	Score	Match	Length	DB	ID	Description
1	900	90.5	1884	3	US-08-753-007A-5	Sequence 5, Appli
2	900	90.5	1884	3	US-09-398-496-5	Sequence 5, Appli
3	881	88.6	993	2	US-08-525-864A-3	Sequence 3, Appli
4	738.6	74.3	3441	2	US-08-525-864A-1	Sequence 1, Appli
5	547.2	55.1	1607	3	US-08-753-007A-3	Sequence 3, Appli
6	547.2	55.1	1607	3	US-09-398-496-3	Sequence 3, Appli
7	492	49.5	1476	3	US-08-753-007A-7	Sequence 7, Appli
8	492	49.5	1476	3	US-09-398-496-7	Sequence 7, Appli
9	492	49.5	2268	3	US-08-753-007A-31	Sequence 31, Appl
10	492	49.5	2268	3	US-09-398-496-31	Sequence 31, Appl
11	467.8	47.1	2467	3	US-08-753-007A-1	Sequence 1, Appli

12	467.8	47.1	2467	3	US-09-398-496-1	Sequence 1, Appli
13	359.6	36.2	1207	2	US-08-525-864A-5	Sequence 5, Appli
14	135.6	13.6	142	2	US-08-525-864A-18	Sequence 18, Appl
15	95.4	9.6	1140	1	US-08-036-555B-149	Sequence 149, App
16	95.4	9.6	1140	1	US-08-469-569-149	Sequence 149, App
17	95.4	9.6	1140	1	US-08-249-322A-149	Sequence 149, App
18	95.4	9.6	1140	1	US-08-469-526A-149	Sequence 149, App
19	95.4	9.6	1140	2	US-08-734-591A-149	Sequence 149, App
20	95.4	9.6	1140	2	US-08-469-660-149	Sequence 149, App
21	95.4	9.6	1140	3	US-08-341-018-55	Sequence 55, Appl
22	95.4	9.6	1140	3	US-08-470-335-149	Sequence 149, App
23	95.4	9.6	1140	3	US-08-735-021-149	Sequence 149, App
24	95.4	9.6	1140	3	US-08-734-664A-149	Sequence 149, App
25	95.4	9.6	1140	3	US-08-470-339-149	Sequence 149, App
26	95.4	9.6	1140	4	US-08-467-602-149	Sequence 149, App
27	95.4	9.6	1140	5	PCT-US94-05083C-145	Sequence 145, App
28	95.4	9.6	1140	5	PCT-US95-06846A-149	Sequence 149, App
29	93.4	9.4	1193	1	US-08-469-526A-134	Sequence 134, App
30	93.4	9.4	1193	2	US-08-734-591A-134	Sequence 134, App
31	93.4	9.4	1193	3	US-08-341-018-3	Sequence 3, Appli
32	93.4	9.4	1193	3	US-08-470-335-134	Sequence 134, App
33	93.4	9.4	1193	3	US-08-735-021-134	Sequence 134, App
34	93.4	9.4	1193	3	US-08-734-664A-134	Sequence 134, App
35	93.4	9.4	1193	3	US-08-470-339-134	Sequence 134, App
36	93.4	9.4	1193	4	US-08-467-602-134	Sequence 134, App
37	91.8	9.2	1193	1	US-08-036-555B-134	Sequence 134, App
38	91.8	9.2	1193	1	US-08-469-569-134	Sequence 134, App
39	91.8	9.2	1193	1	US-08-249-322A-134	Sequence 134, App
40	91.8	9.2	1193	2	US-08-469-660-134	Sequence 134, App
41	91.8	9.2	1193	5	PCT-US94-05083C-130	Sequence 130, App
42	91.8	9.2	1193	5	PCT-US95-06846A-134	Sequence 134, App
43	77.4	7.8	2003	1	US-08-036-555B-21	Sequence 21, Appl
44	77.4	7.8	2003	1	US-08-469-569-21	Sequence 21, Appl
45	77.4	7.8	2003	1	US-08-249-322A-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1

US-08-753-007A-5

; Sequence 5, Application US/08753007A

; Patent No. 6074841

; GENERAL INFORMATION:

; APPLICANT: Gearing, David P.

; APPLICANT: Busfield, Samantha J.

; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES

; TITLE OF INVENTION: AND USES THEREFOR

; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/753,007A
; FILING DATE: 19-NOV-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/699,591
; FILING DATE: 19-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 07334/022001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1884 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 664...1883
; OTHER INFORMATION:
US-08-753-007A-5

```

```

Query Match          90.5%; Score 900; DB 3; Length 1884;
Best Local Similarity 99.3%; Pred. No. 3.5e-223;
Matches 914; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```

```

Qy      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC 60
        |||
Db      218 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC 277

Qy      61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
        |||
Db      278 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 337

Qy      121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
        |||
Db      338 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 397

Qy      181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
        |||
Db      398 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 457

Qy      241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
        |||
Db      458 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 517

Qy      301 CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360

```

```

      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      518 CGCTACATCTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 577
Qy      361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      578 CCCCT-GATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGGC 636
Qy      421 TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      637 TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 696
Qy      481 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 540
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      697 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 756
Qy      541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAAG 600
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      757 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAAG 816
Qy      601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      817 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 876
Qy      661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      877 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 936
Qy      721 ACCACCCTGTCATCCTGGTTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 780
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      937 ACCACCCTGTCATCCTGGTTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 996
Qy      781 TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT 840
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      997 TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT 1056
Qy      841 CCAAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCA 900
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1057 CCAAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCA 1116
Qy      901 GATCCTAAGCAAAGTGTCTCT 920
      |||||||||||| |||
Db      1117 GATCCTAAGCAAAAGCACCT 1136

```

RESULT 2

US-09-398-496-5

; Sequence 5, Application US/09398496

; Patent No. 6133423

; GENERAL INFORMATION:

; APPLICANT: Gearing, David P.

; APPLICANT: Busfield, Samantha J.

; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES

; TITLE OF INVENTION: AND USES THEREFOR

; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

```

; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/398,496
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/753,007
; FILING DATE: 19-NOV-1996
; APPLICATION NUMBER: 08/699,591
; FILING DATE: 19-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 07334/022001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1884 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 664...1883
; OTHER INFORMATION:
US-09-398-496-5

```

```

Query Match          90.5%; Score 900; DB 3; Length 1884;
Best Local Similarity 99.3%; Pred. No. 3.5e-223;
Matches 914; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```

```

Qy      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC 60
          |||
Db      218 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC 277
          |||
Qy      61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
          |||
Db      278 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 337
          |||
Qy      121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
          |||
Db      338 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 397
          |||
Qy      181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
          |||

```

Db	398	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	457
Qy	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	458	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	517
Qy	301	CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	518	CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	577
Qy	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	578	CCCCT-GATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGGC	636
Qy	421	TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Db	637	TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	696
Qy	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	540
Db	697	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	756
Qy	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG	600
Db	757	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG	816
Qy	601	AACTCAGGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Db	817	AACTCAGGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	876
Qy	661	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Db	877	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	936
Qy	721	ACCACCCTGTCATCCTGGTCGGGGCACGCCCAGGAGTGAACGAGACAGCCAAGTCCTAT	780
Db	937	ACCACCCTGTCATCCTGGTCGGGGCACGCCCAGGAGTGAACGAGACAGCCAAGTCCTAT	996
Qy	781	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT	840
Db	997	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT	1056
Qy	841	CCAAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCA	900
Db	1057	CCAAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCA	1116
Qy	901	GATCCTAAGCAAAGTGTCTCT	920
Db	1117	GATCCTAAGCAAAGTGTCTCT	1136

RESULT 3

US-08-525-864A-3

; Sequence 3, Application US/08525864A

; Patent No. 5912326

; GENERAL INFORMATION:

; APPLICANT: Chang, Han

```

; TITLE OF INVENTION: Cerebellum-derived Growth Factors, and Uses
; TITLE OF INVENTION: Related thereto
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: AscII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,864A
; FILING DATE: 8-SEP-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Kara, Catherine J.
; REGISTRATION NUMBER: 41,106
; REFERENCE/DOCKET NUMBER: HUI-017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 993 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..990
US-08-525-864A-3

```

```

Query Match          88.6%; Score 881; DB 2; Length 993;
Best Local Similarity 93.0%; Pred. No. 2.3e-218;
Matches 923; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

```

```

Qy      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCGATGCTGCTCTTCGGTGTGTCACTCGCCTGC 60

Qy      61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TACTCGCCCAGCCTCAAGTCCGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120

Qy      121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
        |||||||||||||||| |||| ||| || |||| |||||||| ||||||||||||
Db      121 GGCAAGGTACAGGGACTGGCCCCGGCAGGCGGTTCCAGCTCTAACAGCACCCGAGAGCCT 180

Qy      181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
        |||||||||||||||||| |||| ||||||||||||||||||||||||||||
Db      181 CCCGCCTCGGGTCGGGTGGCGCTGGTGAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240

```

Qy	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGCGCGCCGCTCGAAAGGAACCAG	300
Qy	301	CGCTACATCTTTTTCCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	301	CGCTACATCTTTTTCCTGGAGCCCACCGAGCAGCCCTTAGTCTTTAAGACAGCCTTTGCC	360
Qy	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	361	CCGGTCGACCCTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Qy	421	TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Db	421	TGCGCAACCCGGCCCCAAGCTGAAGAAGATGAAGAGTCAGACAGGAGAGGTGGGCGAGAAG	480
Qy	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	540
Db	481	CAGTCGCTCAAGTGTGAGGCGGCGGGGAACCCCCAGCCCTCCTATCGATGGTTCAAG	540
Qy	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG	600
Db	541	GACGGCAAGGAGCTCAACCGGAGTCGTGACATTCGCATCAAGTATGGCAACGGCAGAAAG	600
Qy	601	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Db	601	AACTCACGGCTACAGTTCAACAAAGTGAAGGTGGAGGACGCTGGAGAGTACGTCTGTGAG	660
Qy	661	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Db	661	GCTGAGAACATCCTTGGGAAGGACACTGTGAGGGGCCGGCTCCATGTCAACAGTGTGAGC	720
Qy	721	ACCACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	780
Db	721	ACCACTCTGTCGTCTGGTCGGGGCACGCCCGGAAGTGCAATGAGACAGCCAAGTCCTAC	780
Qy	781	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT	840
Db	781	TGTGTGAATGGAGGCGTGTGCTACTACATCGAAGGCATCAACCAACTCTCCTGCAAATGT	840
Qy	841	CCAAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCA	900
Db	841	CCAAACGGATTCTTCGGACAGAGATGTTTGGAGAACTGCC'TTTGCATTGTACATGCCA	900
Qy	901	GATCCTAAGCAAAGTGTCTGTGGGATACACCGGGGACAGGTGTCAGCAGTTCGCAATGG	960
Db	901	GATCCTAAGCAAAGTGTCTGTGGGATACACCGGGGACAGGTGTCAGCAGTTCGCAATGG	960
Qy	961	TCAACTTCTCCAAGCACCTTGGATTTGAATTAA	993
Db	961	TCAACTTCTCCAAGCACCTTGGATTTGAATTAA	993

RESULT 4

US-08-525-864A-1

; Sequence 1, Application US/08525864A


```

; Patent No. 5912326
; GENERAL INFORMATION:
;   APPLICANT:  Chang, Han
;   TITLE OF INVENTION:  Cerebellum-derived Growth Factors, and Uses
;   TITLE OF INVENTION:  Related thereto
;   NUMBER OF SEQUENCES:  18
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  LAHIVE & COCKFIELD
;     STREET:  28 State Street
;     CITY:  Boston
;     STATE:  Massachusetts
;     COUNTRY:  USA
;     ZIP:  02109
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  AscII (text)
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/525,864A
;     FILING DATE:  8-SEP-1995
;     CLASSIFICATION:  530
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Kara, Catherine J.
;     REGISTRATION NUMBER:  41,106
;     REFERENCE/DOCKET NUMBER:  HUI-017
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  (617)227-7400
;     TELEFAX:  (617)742-4214
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  3441 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  180..2441
US-08-525-864A-1

```

```

Query Match          74.3%;  Score 738.6;  DB 2;  Length 3441;
Best Local Similarity 90.4%;  Pred. No. 2.2e-181;
Matches 789;  Conservative 0;  Mismatches 84;  Indels 0;  Gaps 0;

```

```

Qy      1  ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC 60
          |||
Db      180 ATGAGGCGCGACCCGGCCCCCGGCTTCTCGATGCTGCTCTTCGGTGTGTCACTCGCCTGC 239
          |||

Qy      61  TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
          |||
Db      240 TACTCGCCCAGCCTCAAGTCCGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 299
          |||

Qy      121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
          |||
Db      300 GGCAAGGTACAGGGACTGGCCCCGGCAGGCGGTTCCAGCTCTAACAGCACCCGAGAGCCT 359
          |||

```

Qy 181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
 |||||
 Db 360 CCCGCCTCGGGTCGGGTGGCGCTGGTGAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 419
 Qy 241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
 |||||
 Db 420 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGCGCGCCGCTCGAAAGGAACCAG 479
 Qy 301 CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
 |||||
 Db 480 CGCTACATCTTTTTCCTGGAGCCACCGAGCAGCCCTTAGTTTTTAAGACAGCCTTTGCC 539
 Qy 361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
 || |||||
 Db 540 CCGGTCGACCCTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 599
 Qy 421 TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480
 |||||
 Db 600 TGCGCAACCCGGCCCAAGCTGAAGAAGATGAAGAGTCAGACAGGAGAGGTGGGCAGAGAAG 659
 Qy 481 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 540
 || |||||
 Db 660 CAGTCGCTCAAGTGTGAGGCGGCGGCGGGGAACCCAGCCTCCTATCGATGGTTCAAG 719
 Qy 541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTGCGATCAAATATGGCAACGGCAGAAAAG 600
 || |||||
 Db 720 GACGGCAAGGAGCTCAACCGGAGTCGTGACATTGCGATCAAGTATGGCAACGGCAGAAAAG 779
 Qy 601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660
 |||||
 Db 780 AACTCACGGCTACAGTTCAACAAAGTGAAGGTGGAGGACGCTGGAGAGTACGTCTGTGAG 839
 Qy 661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720
 || |||||
 Db 840 GCTGAGAACATCCTTGGGAAGGACACTGTGAGGGGCCGGCTCCATGTCAACAGTGTGAGC 899
 Qy 721 ACCACCCTGTCATCCTGGTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 780
 |||||
 Db 900 ACCACTCTGTCTGCTGCTGGTCGGGGCAGCCCCGGAAGTGCAATGAGACAGCCAAGTCCTAC 959
 Qy 781 TCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT 840
 || |||||
 Db 960 TGTGTGAATGGAGGCGTGTGCTACTACATCGAAGGCATCAACCAACTCTCCTGCAAGTGT 1019
 Qy 841 CCAAATGGATTCTTCGGACAGAGATGTTTGGAG 873
 || |||||
 Db 1020 CCTGTGGGATACACCGGGGACAGGTGTCAGCAG 1052

RESULT 5

US-08-753-007A-3

; Sequence 3, Application US/08753007A

; Patent No. 6074841

; GENERAL INFORMATION:

; APPLICANT: Gearing, David P.

; APPLICANT: Busfield, Samantha J.

; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES

```

; TITLE OF INVENTION:  AND USES THEREFOR
; NUMBER OF SEQUENCES:  33
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE:  Fish & Richardson P.C.
;   STREET:  225 Franklin Street
;   CITY:  Boston
;   STATE:  MA
;   COUNTRY:  US
;   ZIP:  02110-2804
; COMPUTER READABLE FORM:
;   MEDIUM TYPE:  Diskette
;   COMPUTER:  IBM Compatible
;   OPERATING SYSTEM:  DOS
;   SOFTWARE:  FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/08/753,007A
;   FILING DATE:  19-NOV-1996
;   CLASSIFICATION:  536
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:  08/699,591
;   FILING DATE:  19-AUG-1996
; ATTORNEY/AGENT INFORMATION:
;   NAME:  Fasse, J. Peter
;   REGISTRATION NUMBER:  32,983
;   REFERENCE/DOCKET NUMBER:  07334/022001
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE:  617-542-5070
;   TELEFAX:  617-542-8906
;   TELEX:
; INFORMATION FOR SEQ ID NO:  3:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  1607 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  single
;     TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;       NAME/KEY:  Coding Sequence
;       LOCATION:  79...621
;     OTHER INFORMATION:
US-08-753-007A-3

```

```

Query Match          55.1%;  Score 547.2;  DB 3;  Length 1607;
Best Local Similarity 92.3%;  Pred. No. 4.7e-132;
Matches 576;  Conservative  0;  Mismatches  48;  Indels  0;  Gaps  0;

```

```

Qy      371 CCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCC 430
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      2   CTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCC 61

Qy      431 GGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAATCGCTGA 490
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      62  GGCCCAAGCTGAAGAAGATGAAGAGCCAGACAGGAGAGGTGGGTGAGAAGCAGTCGCTCA 121

Qy      491 AGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGATGGCAAGG 550
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      122 AGTGTGAGGCAGCGGCGGGAAACCCCAAGCCCTCCTATCGCTGGTTCAAGGATGGCAAGG 181

```



```

; APPLICATION NUMBER: US/09/398,496
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/753,007
; FILING DATE: 19-NOV-1996
; APPLICATION NUMBER: 08/699,591
; FILING DATE: 19-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 07334/022001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1607 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 79...621
; OTHER INFORMATION:
US-09-398-496-3

```

```

Query Match          55.1%; Score 547.2; DB 3; Length 1607;
Best Local Similarity 92.3%; Pred. No. 4.7e-132;
Matches 576; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

```

```

Qy      371 CCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCC 430
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      2   CTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCC 61

Qy      431 GGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAATCGCTGA 490
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      62  GGCCCAAGCTGAAGAAGATGAAGAGCCAGACAGGAGAGGTGGGTGAGAAGCAGTCGCTCA 121

Qy      491 AGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCTACCGTTGGTTCAAGGATGGCAAGG 550
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      122 AGTGTGAGGCAGCGGCGGGAACCCCCAGCCCTCCTATCGCTGGTTCAAGGATGGCAAGG 181

Qy      551 AGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAACTCACGAC 610
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      182 AACTCAACCGGAGTCGTGATATTCGCATCAAGTATGGCAATGTCAGAAAGAACTCACGGC 241

Qy      611 TACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGCCGAGAACA 670
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      242 TACAGTTCAACAAAGTGAGGGTGGAGGATGCCGGGGAGTACGTCTGTGAGGCCGAGAACA 301

Qy      671 TCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCACCACCCTGT 730
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      302 TCCTTGGGAAGGACACCGTGAGGGGCCGACTCCATGTCAACAGCGTGAGCACCACCTCTGT 361

```



```

;       TELEX:
;  INFORMATION FOR SEQ ID NO: 7:
;  SEQUENCE CHARACTERISTICS:
;    LENGTH: 1476 base pairs
;    TYPE: nucleic acid
;    STRANDEDNESS: single
;    TOPOLOGY: linear
;    MOLECULE TYPE: cDNA
;    FEATURE:
;      NAME/KEY: Coding Sequence
;      LOCATION: 69...1475
;      OTHER INFORMATION:
US-08-753-007A-7

```

```

Query Match          49.5%; Score 492; DB 3; Length 1476;
Best Local Similarity 92.8%; Pred. No. 8.3e-118;
Matches 516; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

```

```

Qy      363 CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG 422
        || || | || | | | | || | | | | | | | |
Db      98 CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCGATCCCGGGGAGAA 157

Qy      423 CGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 482
        |||||||||||||||||||||||||||||||||||||||
Db      158 AGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 217

Qy      483 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 542
        |||||||||||||||||||||||||||||||||||||||
Db      218 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 277

Qy      543 TGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAA 602
        |||||||||||||||||||||||||||||||||||||||
Db      278 TGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAA 337

Qy      603 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 662
        |||||||||||||||||||||||||||||||||||||||
Db      338 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 397

Qy      663 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722
        |||||||||||||||||||||||||||||||||||||||
Db      398 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 457

Qy      723 CACCCTGTCATCCTGGTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 782
        |||||||||||||||||||||||||||||||||||||||
Db      458 CACCCTGTCATCCTGGTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 517

Qy      783 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC 842
        |||||||||||||||||||||||||||||||||||||||
Db      518 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC 577

Qy      843 AAATGGATTCTTCGGACAGAGATGTTTGGAGAAACTGCCTTTGCGATTGTACATGCCAGA 902
        |||||||||||||||||||||||||||||||||||||||
Db      578 AAATGGATTCTTCGGACAGAGATGTTTGGAGAAACTGCCTTTGCGATTGTACATGCCAGA 637

Qy      903 TCCTAAGCAAAGTGTC 918
        ||||||||| | |
Db      638 TCCTAAGCAAAAAGCC 653

```

RESULT 8
 US-09-398-496-7
 ; Sequence 7, Application US/09398496
 ; Patent No. 6133423
 ; GENERAL INFORMATION:
 ; APPLICANT: Gearing, David P.
 ; APPLICANT: Busfield, Samantha J.
 ; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES
 ; TITLE OF INVENTION: AND USES THEREFOR
 ; NUMBER OF SEQUENCES: 33
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson P.C.
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: US
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSEQ Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/398,496
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/753,007
 ; FILING DATE: 19-NOV-1996
 ; APPLICATION NUMBER: 08/699,591
 ; FILING DATE: 19-AUG-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fasse, J. Peter
 ; REGISTRATION NUMBER: 32,983
 ; REFERENCE/DOCKET NUMBER: 07334/022001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617-542-5070
 ; TELEFAX: 617-542-8906
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1476 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: Coding Sequence
 ; LOCATION: 69...1475
 ; OTHER INFORMATION:
 US-09-398-496-7

Query Match 49.5%; Score 492; DB 3; Length 1476;
 Best Local Similarity 92.8%; Pred. No. 8.3e-118;
 Matches 516; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Qy 363 CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG 422
 || || | || | | | | || | | | | | | |
 Db 98 CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCGATCCCGGGGAGAA 157

Qy 423 CGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 482
 |||||
 Db 158 AGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 217

Qy 483 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 542
 |||||
 Db 218 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 277

Qy 543 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 602
 |||||
 Db 278 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 337

Qy 603 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 662
 |||||
 Db 338 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 397

Qy 663 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722
 |||||
 Db 398 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 457

Qy 723 CACCCTGTCATCCTGGTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 782
 |||||
 Db 458 CACCCTGTCATCCTGGTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 517

Qy 783 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACGAGCTCTCCTGCAAATGTCC 842
 |||||
 Db 518 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACGAGCTCTCCTGCAAATGTCC 577

Qy 843 AAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCAGA 902
 |||||
 Db 578 AAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCAGA 637

Qy 903 TCCTAAGCAAAGTGTC 918
 ||||| | |
 Db 638 TCCTAAGCAAAAAGCC 653

RESULT 9

US-08-753-007A-31

; Sequence 31, Application US/08753007A

; Patent No. 6074841

; GENERAL INFORMATION:

; APPLICANT: Gearing, David P.

; APPLICANT: Busfield, Samantha J.

; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES

; TITLE OF INVENTION: AND USES THEREFOR

; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

```

; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/753,007A
; FILING DATE: 19-NOV-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/699,591
; FILING DATE: 19-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 07334/022001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX:
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2268 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 69...2009
; OTHER INFORMATION:
US-08-753-007A-31

```

```

Query Match          49.5%; Score 492; DB 3; Length 2268;
Best Local Similarity 92.8%; Pred. No. 9.7e-118;
Matches 516; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

```

```

Qy      363 CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG 422
      || || | || | | | | || | | | | | | | |
Db      98 CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCGATCCCGGGGAGAA 157

Qy      423 CGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 482
      |||||
Db      158 AGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 217

Qy      483 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 542
      |||||
Db      218 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 277

Qy      543 TGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAA 602
      |||||
Db      278 TGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAA 337

Qy      603 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 662
      |||||

```

Db 338 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 397
 Qy 663 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 398 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 457
 Qy 723 CACCCTGTCATCCTGGTCGGGGCACGCCCAGGAAGTGCAACGAGACAGCCAAGTCCTATTG 782
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 458 CACCCTGTCATCCTGGTCGGGGCACGCCCAGGAAGTGCAACGAGACAGCCAAGTCCTATTG 517
 Qy 783 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC 842
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 518 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC 577
 Qy 843 AAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCAGA 902
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 578 AAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCAGA 637
 Qy 903 TCCTAAGCAAAGTGTC 918
 ||||||||| ||
 Db 638 TCCTAAGCAAAAAGCC 653

RESULT 10

US-09-398-496-31

; Sequence 31, Application US/09398496

; Patent No. 6133423

; GENERAL INFORMATION:

; APPLICANT: Gearing, David P.

; APPLICANT: Busfield, Samantha J.

; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES

; TITLE OF INVENTION: AND USES THEREFOR

; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/398,496

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/753,007

; FILING DATE: 19-NOV-1996

; APPLICATION NUMBER: 08/699,591

; FILING DATE: 19-AUG-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Fasse, J. Peter

; REGISTRATION NUMBER: 32,983

```

; REFERENCE/DOCKET NUMBER: 07334/022001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX:
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2268 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 69...2009
; OTHER INFORMATION:
US-09-398-496-31

```

```

Query Match          49.5%; Score 492; DB 3; Length 2268;
Best Local Similarity 92.8%; Pred. No. 9.7e-118;
Matches 516; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

```

```

Qy      363 CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG 422
          || || | || | || | || | || | || | || | || |
Db      98  CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCGATCCCGGGGAGAA 157

Qy      423 CGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 482
          |||||
Db      158 AGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 217

Qy      483 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 542
          |||||
Db      218 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 277

Qy      543 TGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAA 602
          |||||
Db      278 TGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAA 337

Qy      603 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 662
          |||||
Db      338 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 397

Qy      663 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722
          |||||
Db      398 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 457

Qy      723 CACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 782
          |||||
Db      458 CACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 517

Qy      783 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC 842
          |||||
Db      518 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC 577

Qy      843 AAATGGATTCTTCGGACAGAGATGTTTGGAGAAACTGCCTTTGCGATTGTACATGCCAGA 902
          |||||
Db      578 AAATGGATTCTTCGGACAGAGATGTTTGGAGAAACTGCCTTTGCGATTGTACATGCCAGA 637

```

Qy 903 TCCTAAGCAAAGTGC 918
| | | | | | | | | | | |
Db 638 TCCTAAGCAAAAAGCC 653

RESULT 11

US-08-753-007A-1

; Sequence 1, Application US/08753007A

; Patent No. 6074841

; GENERAL INFORMATION:

; APPLICANT: Gearing, David P.

; APPLICANT: Busfield, Samantha J.

; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES

; TITLE OF INVENTION: AND USES THEREFOR

; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/753,007A

; FILING DATE: 19-NOV-1996

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/699,591

; FILING DATE: 19-AUG-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Fasse, J. Peter

; REGISTRATION NUMBER: 32,983

; REFERENCE/DOCKET NUMBER: 07334/022001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-542-5070

; TELEFAX: 617-542-8906

; TELEX:

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2467 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: circular

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: Coding Sequence

; LOCATION: 79...1893

; OTHER INFORMATION:

US-08-753-007A-1

Query Match

47.1%; Score 467.8; DB 3; Length 2467;


```

;   ADDRESSEE: LAHIVE & COCKFIELD
;   STREET: 28 State Street
;   CITY: Boston
;   STATE: Massachusetts
;   COUNTRY: USA
;   ZIP: 02109
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: AscII (text)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/525,864A
;   FILING DATE: 8-SEP-1995
;   CLASSIFICATION: 530
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Kara, Catherine J.
;   REGISTRATION NUMBER: 41,106
;   REFERENCE/DOCKET NUMBER: HUI-017
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (617)227-7400
;   TELEFAX: (617)742-4214
;   INFORMATION FOR SEQ ID NO: 18:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 142 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
US-08-525-864A-18

```

```

Query Match          13.6%; Score 135.6; DB 2; Length 142;
Best Local Similarity 97.2%; Pred. No. 4.2e-26;
Matches 138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

Qy      792 AGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCCAAATGGATT 851
          ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      1 AGGCGTGTGCTACTACATCGAAGGCATCAACCAACTCTCCTGCAAATGTCCAAACGGATT 60

Qy      852 CTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCAGATCCTAAGCA 911
          ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      61 CTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCAGATCCTAAGCA 120

Qy      912 AAGTGTCTGTGGGATACACCG 933
          ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      121 AAGTGTCTGTGGGATACACCG 142

```

RESULT 15

US-08-036-555B-149

; Sequence 149, Application US/08036555B

; Patent No. 5530109

; GENERAL INFORMATION:

; APPLICANT: Goodearl, Andrew; Stroobant, Paul;

; APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;

; APPLICANT: Chen, Maio Su; Hiles, Ian

; TITLE OF INVENTION: Glial Mitogenic Factors, Their

Db	71	GGGGCCACCCGCTTCCCCTCCTGCGGGCGCCTCAAGGAGGACAGCAGGTACATCTTCT	130
Qy	314	TCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCCCCCTCGATACC-	372
Db	131	TCATGGAGCCCGAGGCCAACAGCAGCGGCGGGCCCGGCCGCTTCCGAGCCTCCTTCCCC	190
Qy	373	-----AACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTGC	423
Db	191	CCTCTCGAGACGGGCGGAACCTCAAGAAGGAGGTGAGCCGGGTGCTGTGCAACGG-TGC	249
Qy	424	GCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAA	483
Db	250	GCCTTGCCTCCCCGCTTGAAAGAGATGAAGAGTCAGGAGTCTGTGGCAGGTTCCAAACTA	309
Qy	484	TCGCTGAAGTGTGAGGCAGCAGCCGGAATCCCCAGCCTTCCCTACCGTTGGTTCAAGGAT	543
Db	310	GTGCTTCGGTGCAGACCAAGTCTGAATACTCCTCTCTCAAGTTCAAGTGGTTCAAGAAT	369
Qy	544	GGCAAGGAGCTCAACCG---CAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG	600
Db	370	GGGAGTGAATTAAGCCGAAAGAACAAACCAGAAAACATCAAGATACAGAAAAGGCCGGGG	429
Qy	601	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Db	430	AAGTCAGAACTTCGCATTAGCAAAGCGTCACTGGCTGATTCTGGAGAATATATGTGCAAA	489
Qy	661	GCCGAGAACATCCTGGGGAAGGACA-----CCGTCCGGGGCGGCTTTACGTCAACAGC	714
Db	490	GTGATCAGCAAACTAGGAAATGACAGTGCCTCTGCCAACATCACCATTGTGGAGTCAAAC	549
Qy	715	GTGAGCACCACCCTGTCATCCTGGTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAG	774
Db	550	GCCACATCCACATCTACAGCTGGGACAAGCCATCTTGTCAAGTGTGCAGAGAAGGAGAAA	609
Qy	775	TCCTATTGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTC----	830
Db	610	ACTTTCTGTGTGAATGGAGGCGAGTGCTTCATGGTGAAAGACCTTTCAAATCCCTCAAGA	669
Qy	831	-----CTGCAAATGTCCAAATGGATTCTTCGGACAGAGATGTTTGGAGAAACTGCCTTTG	885
Db	670	TACTTGTGCAAGTGCCAACCTGGATTCACTGGAGCGAGATGTACTGAGAATGTGCCCATG	729
Qy	886	CGATTGTACATGCCAGATCCTAAGCAAAGTGTCTGTGGGATACACCGGGGACAGGTGTC	945
Db	730	AAAGTCCAAA-----CCCAAGAAAGTGCCCAAATGAGTTTACTGGTGATCGCTGCC	781
Qy	946	AGCAGTTCGCAATGGTCAACTTCTCCA	972
Db	782	AAACTACGTAATGGCCAGCTTCTACA	808

Search completed: January 14, 2004, 10:25:23
 Job time : 78.5907 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2004, 10:23:13 ; Search time 384.249 Seconds
(without alignments)
9118.095 Million cell updates/sec

Title: US-09-864-675-1
Perfect score: 994
Sequence: 1 atgaggcgcgacccggcccc.....caccttgatttgaattaaa 994

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2324096 seqs, 1762381658 residues

Total number of hits satisfying chosen parameters: 4648192

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8

Result	Query					
No.	Score	Match	Length	DB	ID	Description

1	994	100.0	994	9	US-09-864-675-1	Sequence 1, Appli
2	900	90.5	1884	14	US-10-096-241-5	Sequence 5, Appli
3	849	85.4	897	9	US-09-864-675-3	Sequence 3, Appli
4	547.2	55.1	1607	14	US-10-096-241-3	Sequence 3, Appli
5	492	49.5	1476	14	US-10-096-241-7	Sequence 7, Appli
6	492	49.5	2268	14	US-10-096-241-31	Sequence 31, Appl
7	467.8	47.1	2467	14	US-10-096-241-1	Sequence 1, Appli
8	173	17.4	201	13	US-10-029-386-26613	Sequence 26613, A
9	173	17.4	573	13	US-10-029-386-12913	Sequence 12913, A
c 10	113.6	11.4	579	13	US-10-029-386-2532	Sequence 2532, Ap
c 11	111.8	11.2	171	13	US-10-029-386-16232	Sequence 16232, A
12	95.4	9.6	1140	7	US-08-736-019-149	Sequence 149, App
13	95.4	9.6	1140	11	US-09-366-886-55	Sequence 55, Appl
14	93.4	9.4	1193	7	US-08-736-019-134	Sequence 134, App
15	93.4	9.4	1193	11	US-09-366-886-3	Sequence 3, Appli
16	77.4	7.8	1986	11	US-09-373-658-71	Sequence 71, Appl
17	77.4	7.8	1986	12	US-09-989-687-71	Sequence 71, Appl
18	77.4	7.8	2003	7	US-08-736-019-21	Sequence 21, Appl
19	77.4	7.8	2003	11	US-09-366-886-71	Sequence 71, Appl
20	77.4	7.8	2003	11	US-09-373-658-72	Sequence 72, Appl
21	77.4	7.8	2003	12	US-09-989-687-72	Sequence 72, Appl
22	76.8	7.7	1108	7	US-08-736-019-135	Sequence 135, App
23	76.8	7.7	1108	11	US-09-366-886-5	Sequence 5, Appli
24	72.8	7.3	744	7	US-08-736-019-133	Sequence 133, App
25	72.8	7.3	744	11	US-09-366-886-1	Sequence 1, Appli
26	60.4	6.1	848	13	US-10-027-632-158930	Sequence 158930,
27	60.4	6.1	848	14	US-10-027-632-158930	Sequence 158930,
28	52	5.2	1092	15	US-10-156-761-4747	Sequence 4747, Ap
29	52	5.2	9025608	15	US-10-156-761-1	Sequence 1, Appli
30	50.6	5.1	3111	9	US-09-773-517-12	Sequence 12, Appl
31	50.6	5.1	3111	9	US-09-792-025-12	Sequence 12, Appl
32	50.6	5.1	3111	9	US-09-849-868-12	Sequence 12, Appl
33	50.6	5.1	3111	13	US-10-453-183-12	Sequence 12, Appl
34	50.6	5.1	3111	15	US-10-290-578-1	Sequence 1, Appli
35	49.8	5.0	1935	10	US-09-778-510-21	Sequence 21, Appl
36	49.8	5.0	1935	10	US-09-778-187B-3	Sequence 3, Appli
37	49.8	5.0	1935	13	US-10-302-041-21	Sequence 21, Appl
38	46.8	4.7	1652	7	US-08-736-019-148	Sequence 148, App
39	46.8	4.7	1652	11	US-09-366-886-53	Sequence 53, Appl
40	46	4.6	1521	15	US-10-156-761-5445	Sequence 5445, Ap
41	45.8	4.6	543	11	US-09-918-995-26620	Sequence 26620, A
42	45.4	4.6	1092	15	US-10-156-761-6888	Sequence 6888, Ap
43	44.6	4.5	477	11	US-09-918-995-28888	Sequence 28888, A
44	44.6	4.5	1700	14	US-10-071-766-31	Sequence 31, Appl
45	44.6	4.5	1715	9	US-09-773-517-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1

US-09-864-675-1

; Sequence 1, Application US/09864675

; Patent No. US20020081286A1

; GENERAL INFORMATION:

; APPLICANT: Marchionni, Mark

```
; TITLE OF INVENTION: NRG-2 NUCLEIC ACID MOLECULES,
; TITLE OF INVENTION: POLYPEPTIDES, AND DIAGNOSTIC AND THERAPEUTIC METHODS
; FILE REFERENCE: 04585/049002
; CURRENT APPLICATION NUMBER: US/09/864,675
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/206,495
; PRIOR FILING DATE: 2000-05-23
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 994
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-864-675-1
```

```
Query Match          100.0%; Score 994; DB 9; Length 994;
Best Local Similarity 100.0%; Pred. No. 1.3e-279;
Matches 994; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC 60
        |||
Db      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC 60

Qy     61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
        |||
Db     61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120

Qy    121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
        |||
Db    121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180

Qy    181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
        |||
Db    181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240

Qy    241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
        |||
Db    241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300

Qy    301 CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
        |||
Db    301 CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360

Qy    361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
        |||
Db    361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420

Qy    421 TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480
        |||
Db    421 TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480

Qy    481 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCAGCCTTCTACCGTTGGTTCAAG 540
        |||
Db    481 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCAGCCTTCTACCGTTGGTTCAAG 540

Qy    541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTGCGATCAAATATGGCAACGGCAGAAAAG 600
        |||
```

Db	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG	600
Qy	601	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Db	601	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Qy	661	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Db	661	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Qy	721	ACCACCCTGTCATCCTGGTCGGGGCACGCCCGBAAGTGCAACGAGACAGCCAAGTCCTAT	780
Db	721	ACCACCCTGTCATCCTGGTCGGGGCACGCCCGBAAGTGCAACGAGACAGCCAAGTCCTAT	780
Qy	781	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT	840
Db	781	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT	840
Qy	841	CCAAATGGATTCTTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCA	900
Db	841	CCAAATGGATTCTTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCA	900
Qy	901	GATCCTAAGCAAAGTGTCTGTGGGATACACCGGGGACAGGTGTCAGCAGTTCGCAATGG	960
Db	901	GATCCTAAGCAAAGTGTCTGTGGGATACACCGGGGACAGGTGTCAGCAGTTCGCAATGG	960
Qy	961	TCAACTTCTCCAAGCACCTTGATTTGAATTAA	994
Db	961	TCAACTTCTCCAAGCACCTTGATTTGAATTAA	994

RESULT 2

US-10-096-241-5

; Sequence 5, Application US/10096241

; Publication No. US20020127594A1

; GENERAL INFORMATION:

; APPLICANT: Gearing, David P.

; Busfield, Samantha J.

; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES
; AND USES THEREFOR

; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/096,241

; FILING DATE: 12-Mar-2002

; CLASSIFICATION: <Unknown>


```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/699,591
; FILING DATE: 19-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 07334/022001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1884 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 664...1883
; OTHER INFORMATION:
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-096-241-5

```

```

Query Match          90.5%; Score 900; DB 14; Length 1884;
Best Local Similarity 99.3%; Pred. No. 4.1e-252;
Matches 914; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```

```

Qy      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCGCTCGCCTGC 60
        |||
Db      218 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCGCTCGCCTGC 277
Qy      61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
        |||
Db      278 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 337
Qy      121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
        |||
Db      338 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 397
Qy      181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
        |||
Db      398 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 457
Qy      241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
        |||
Db      458 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 517
Qy      301 CGCTACATCTTTTTCCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
        |||
Db      518 CGCTACATCTTTTTCCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 577
Qy      361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
        |||
Db      578 CCCCT-GATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGGC 636

```

Qy 421 TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480
 |||
 Db 637 TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 696
 Qy 481 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 540
 |||
 Db 697 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 756
 Qy 541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAAG 600
 |||
 Db 757 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAAG 816
 Qy 601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660
 |||
 Db 817 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 876
 Qy 661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720
 |||
 Db 877 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 936
 Qy 721 ACCACCCTGTCATCCTGGTCCGGGACAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 780
 |||
 Db 937 ACCACCCTGTCATCCTGGTCCGGGACAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 996
 Qy 781 TCGGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT 840
 |||
 Db 997 TCGGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT 1056
 Qy 841 CCAAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCA 900
 |||
 Db 1057 CCAAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCA 1116
 Qy 901 GATCCTAAGCAAAGTGTCT 920
 |||
 Db 1117 GATCCTAAGCAAAGCACCT 1136

RESULT 3

US-09-864-675-3

; Sequence 3, Application US/09864675

; Patent No. US20020081286A1

; GENERAL INFORMATION:

; APPLICANT: Marchionni, Mark

; TITLE OF INVENTION: NRG-2 NUCLEIC ACID MOLECULES,

; TITLE OF INVENTION: POLYPEPTIDES, AND DIAGNOSTIC AND THERAPEUTIC METHODS

; FILE REFERENCE: 04585/049002

; CURRENT APPLICATION NUMBER: US/09/864,675

; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/206,495

; PRIOR FILING DATE: 2000-05-23

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 897

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-864-675-3

Query Match 85.4%; Score 849; DB 9; Length 897;
 Best Local Similarity 98.3%; Pred. No. 2.4e-237;
 Matches 858; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC	60
Db	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC	60
Qy	61	TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Db	61	TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Qy	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Db	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Qy	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Db	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Qy	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Qy	301	CGCTACATCTTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	301	CGCTACATCTTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Qy	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Qy	421	TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Db	421	TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Qy	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCAGCCTTCCTACCGTTGGTTCAAG	540
Db	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCAGCCTTCCTACCGTTGGTTCAAG	540
Qy	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG	600
Db	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG	600
Qy	601	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Db	601	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Qy	661	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Db	661	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Qy	721	ACCACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	780
Db	721	ACCACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	780

Qy 781 TCGGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT 840
 |||||
 Db 781 TCGGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT 840
 Qy 841 CCAAATGGATTCTTCGGACAGAGATGTTTGGAG 873
 || |||| | ||| | || ||| | ||
 Db 841 CCTGTGGGATACACCGGGGACAGGTGTCAGCAG 873

RESULT 4

US-10-096-241-3

; Sequence 3, Application US/10096241

; Publication No. US20020127594A1

; GENERAL INFORMATION:

; APPLICANT: Gearing, David P.

; Busfield, Samantha J.

; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES
 ; AND USES THEREFOR

; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/096,241

; FILING DATE: 12-Mar-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/699,591

; FILING DATE: 19-AUG-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Fasse, J. Peter

; REGISTRATION NUMBER: 32,983

; REFERENCE/DOCKET NUMBER: 07334/022001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-542-5070

; TELEFAX: 617-542-8906

; TELEX: <Unknown>

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1607 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: Coding Sequence

; LOCATION: 79...621

; OTHER INFORMATION:


```

; APPLICANT: Gearing, David P.
;           Busfield, Samantha J.
; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES
;                   AND USES THEREFOR
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
;           ADDRESSEE: Fish & Richardson P.C.
;           STREET: 225 Franklin Street
;           CITY: Boston
;           STATE: MA
;           COUNTRY: US
;           ZIP: 02110-2804
; COMPUTER READABLE FORM:
;           MEDIUM TYPE: Diskette
;           COMPUTER: IBM Compatible
;           OPERATING SYSTEM: DOS
;           SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
;           APPLICATION NUMBER: US/10/096,241
;           FILING DATE: 12-Mar-2002
;           CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
;           APPLICATION NUMBER: 08/699,591
;           FILING DATE: 19-AUG-1996
; ATTORNEY/AGENT INFORMATION:
;           NAME: Fasse, J. Peter
;           REGISTRATION NUMBER: 32,983
;           REFERENCE/DOCKET NUMBER: 07334/022001
; TELECOMMUNICATION INFORMATION:
;           TELEPHONE: 617-542-5070
;           TELEFAX: 617-542-8906
;           TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 7:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1476 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   FEATURE:
;     NAME/KEY: Coding Sequence
;     LOCATION: 69...1475
;     OTHER INFORMATION:
;   SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-096-241-7

```

```

Query Match          49.5%; Score 492; DB 14; Length 1476;
Best Local Similarity 92.8%; Pred. No. 4e-133;
Matches 516; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

```

```

Qy      363 CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG 422
          || || | || | | | | || | || | | || | | ||
Db      98  CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCGATCCCGGGGAGAA 157

Qy      423 CGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 482
          |||||
Db      158 AGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 217

```

Qy 483 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 542
 |||
 Db 218 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 277
 Qy 543 TGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAA 602
 |||
 Db 278 TGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAA 337
 Qy 603 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 662
 |||
 Db 338 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 397
 Qy 663 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722
 |||
 Db 398 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 457
 Qy 723 CACCCTGTCATCCTGGTTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 782
 |||
 Db 458 CACCCTGTCATCCTGGTTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 517
 Qy 783 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC 842
 |||
 Db 518 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC 577
 Qy 843 AAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCAGA 902
 |||
 Db 578 AAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCAGA 637
 Qy 903 TCCTAAGCAAAGTGTC 918
 |||
 Db 638 TCCTAAGCAAAAAGCC 653

RESULT 6

US-10-096-241-31

; Sequence 31, Application US/10096241

; Publication No. US20020127594A1

; GENERAL INFORMATION:

; APPLICANT: Gearing, David P.

; Busfield, Samantha J.

; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES
 AND USES THEREFOR

; NUMBER OF SEQUENCES: 33.

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ Version 2.0

; CURRENT APPLICATION DATA:

```

; APPLICATION NUMBER: US/10/096,241
; FILING DATE: 12-Mar-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/699,591
; FILING DATE: 19-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 07334/022001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2268 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 69...2009
; OTHER INFORMATION:
; SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-10-096-241-31

```

```

Query Match          49.5%; Score 492; DB 14; Length 2268;
Best Local Similarity 92.8%; Pred. No. 4.6e-133;
Matches 516; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

```

```

Qy      363 CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG 422
      || || | || | | | | || | | | | | | | |
Db      98 CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCGATCCCGGGGAGAA 157

Qy      423 CGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 482
      |||||
Db      158 AGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 217

Qy      483 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 542
      |||||
Db      218 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 277

Qy      543 TGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAA 602
      |||||
Db      278 TGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAA 337

Qy      603 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 662
      |||||
Db      338 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 397

Qy      663 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722
      |||||
Db      398 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 457

Qy      723 CACCCTGTCATCCTGGTTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 782

```



```

      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      458 CACCCTGTCATCCTGGTCGGGGCACGCCCAGGAAGTGCAACGAGACAGCCAAGTCCTATTG 517
Qy      783 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC 842
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      518 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC 577
Qy      843 AAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCAGA 902
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      578 AAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCAGA 637
Qy      903 TCCTAAGCAAAGTGTC 918
      |||||||||  | |
Db      638 TCCTAAGCAAAAAGCC 653

```

RESULT 7

US-10-096-241-1

; Sequence 1, Application US/10096241

; Publication No. US20020127594A1

; GENERAL INFORMATION:

; APPLICANT: Gearing, David P.

; Busfield, Samantha J.

; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES

; AND USES THEREFOR

; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/096,241

; FILING DATE: 12-Mar-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/699,591

; FILING DATE: 19-AUG-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Fasse, J. Peter

; REGISTRATION NUMBER: 32,983

; REFERENCE/DOCKET NUMBER: 07334/022001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-542-5070

; TELEFAX: 617-542-8906

; TELEX: <Unknown>

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2467 base pairs

; TYPE: nucleic acid

```

;          STRANDEDNESS: single
;          TOPOLOGY: circular
;          MOLECULE TYPE: cDNA
;          FEATURE:
;          NAME/KEY: Coding Sequence
;          LOCATION: 79...1893
;          OTHER INFORMATION:
;          SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-096-241-1

```

```

Query Match          47.1%; Score 467.8; DB 14; Length 2467;
Best Local Similarity 91.3%; Pred. No. 5.5e-126;
Matches 496; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

```

```

Qy      371 CCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCC 430
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      2   CTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCC 61

Qy      431 GGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAATCGCTGA 490
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      62 GGCCCCAAGCTGAAGAAGATGAAGAGCCAGACAGGAGAGGTGGGTGAGAAGCAGTCGCTCA 121

Qy      491 AGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGATGGCAAGG 550
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      122 AGTGTGAGGCAGCGGCGGGAACCCCCAGCCCTCCTATCGCTGGTTCAAGGATGGCAAGG 181

Qy      551 AGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAACTCACGAC 610
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      182 AACTCAACCGGAGTCGTGATATTTCGCATCAAGTATGGCAATGTCAGAAAGAACTCACGGC 241

Qy      611 TACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGCCGAGAACA 670
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      242 TACAGTTCAACAAAGTGAGGGTGGAGGATGCCGGGGAGTACGTCTGTGAGGCCGAGAACA 301

Qy      671 TCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCACCACCCTGT 730
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      302 TCCTTGGGAAGGACACCGTGAGGGGCGGACTCCATGTCAACAGCGTGAGCACCACCTCTGT 361

Qy      731 CATCCTGGTCGGGGCACGCCCAGGAGTGAACGAGACAGCCAAGTCCTATTGCGTCAATG 790
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      362 CATCCTGGTCGGGACATGCCCAGGAGTGAATGAGACCGCCAAGTCCTACTGTGTGAATG 421

Qy      791 GAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCCAAATGGAT 850
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      422 GAGGCGTGTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCCAAACGGAT 481

Qy      851 TCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCAGATCCTAAGC 910
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      482 TCTTCGGACAGAGATGTTTGGAGAACTGCCTTTGCGATTGTACATGCCAGATCCTAAGC 541

Qy      911 AAA 913
          | | |
Db      542 AAA 544

```

```

RESULT 8
US-10-029-386-26613

```

```

; Sequence 26613, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES
USEFUL FOR GENE
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 26613
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR5.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.55
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.49
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.72
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.66
; OTHER INFORMATION: SWISSPROT HIT: O14511, EVALUE 3.00e-29
; OTHER INFORMATION: NT HIT: AF119152.1, EVALUE 1.00e-109
; OTHER INFORMATION: EST_HUMAN HIT: BF108794.1, EVALUE 3.00e-93
US-10-029-386-26613

```

```

Query Match          17.4%; Score 173; DB 13; Length 201;
Best Local Similarity 100.0%; Pred. No. 2.5e-40;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      424 GCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAA 483
          |||
Db      27  GCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAA 86

Qy      484 TCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGAT 543
          |||
Db      87  TCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGAT 146

Qy      544 GGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAG 596
          |||
Db      147 GGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAG 199

```

RESULT 9

```

US-10-029-386-12913
; Sequence 12913, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES
USEFUL FOR GENE
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO

```

```
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 12913
; LENGTH: 573
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR5.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.55
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.49
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.72
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.66
; OTHER INFORMATION: SWISSPROT HIT: O14511, EVALUE 2.00e-28
; OTHER INFORMATION: NT HIT: AF119152.1, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: BG996653.1, EVALUE 1.00e-108
US-10-029-386-12913
```

```
Query Match          17.4%; Score 173; DB 13; Length 573;
Best Local Similarity 100.0%; Pred. No. 3.4e-40;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      424 GCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAA 483
          |||
Db      377 GCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAA 436

Qy      484 TCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGAT 543
          |||
Db      437 TCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGAT 496

Qy      544 GGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAG 596
          |||
Db      497 GGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAG 549
```

RESULT 10

US-10-029-386-2532/c

```
; Sequence 2532, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES
USEFUL FOR GENE
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 2532
; LENGTH: 579
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
; FEATURE:
; OTHER INFORMATION: MAP TO CHR5.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.6
; OTHER INFORMATION: NT HIT: AF119153.1, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: BF108794.1, EVALUE 2.00e-57
; OTHER INFORMATION: SWISSPROT HIT: O14511, EVALUE 2.00e-12
US-10-029-386-2532
```

```
Query Match          11.4%; Score 113.6; DB 13; Length 579;
Best Local Similarity 92.9%; Pred. No. 7.4e-23;
Matches 130; Conservative 0; Mismatches 9; Indels 1; Gaps 1;
```

```
Qy      594 CAGAAAGAACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGT 653
          |||
Db      489 CAGAAAGAACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGT 430
          |||

Qy      654 CTGCGAGGCCGAGAACATCCTGGGGAAGGACACCG-TCCGGGGCCGGCTTTACGTCAACA 712
          |||
Db      429 CTGCGAGGCCGAGAACATCCTGGGGAAGGACACCGCTCCGGGGCCGGCTTTACGTCAACA 370
          |||

Qy      713 GCGTGAGCACCAACCCTGTCA 732
          |||
Db      369 GCGGTAGGTGGGCCAGACA 350
```

RESULT 11

```
US-10-029-386-16232/c
; Sequence 16232, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES
USEFUL FOR GENE
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 16232
; LENGTH: 171
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR5.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.3
```

; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.2
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.4
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.6
 ; OTHER INFORMATION: NT HIT: AF119153.1, EVALUE 7.00e-87
 ; OTHER INFORMATION: SWISSPROT HIT: O14511, EVALUE 3.00e-13
 ; OTHER INFORMATION: EST_HUMAN HIT: BF108794.1, EVALUE 5.00e-58
 US-10-029-386-16232

Query Match 11.2%; Score 111.8; DB 13; Length 171;
 Best Local Similarity 97.6%; Pred. No. 1.7e-22;
 Matches 124; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 594 CAGAAAGAACTCAGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGT 653
 |||
 Db 130 CAGAAAGAACTCAGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGT 71
 Qy 654 CTGCGAGGCCGAGAACATCCTGGGGAAGGACACCG-TCCGGGGCCGGCTTTACGTCAACA 712
 |||
 Db 70 CTGCGAGGCCGAGAACATCCTGGGGAAGGACACCGCTCCGGGGCCGGCTTTACGTCAACA 11
 Qy 713 GCGTGAG 719
 |||
 Db 10 GCGGTAG 4

RESULT 12

US-08-736-019-149

; Sequence 149, Application US/08736019

; Publication No. US20030207799A1

; GENERAL INFORMATION:

; APPLICANT: Goodearl, Andrew

; APPLICANT: Stroobant, Paul

; APPLICANT: Minghetti, Luisa

; APPLICANT: Waterfield, Michael

; APPLICANT: Marchionni, Mark

; APPLICANT: Chen, Mario

; APPLICANT: Hiles, Ian

; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR

; TITLE OF INVENTION: PREPARATION AND USE

; NUMBER OF SEQUENCES: 189

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Clark & Elbing LLP

; STREET: 176 Federal Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: U.S.A.

; ZIP: 02110

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

; COMPUTER: IBM Compatible Pentium

; OPERATING SYSTEM: Windows95

; SOFTWARE: FastSeq Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/736,019

; FILING DATE: 22-OCT-1996

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: 08/471,833
; FILING DATE: 06-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/036,555
; FILING DATE: 24-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/965,173
; FILING DATE: 23-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/907,138
; FILING DATE: 30-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/940,389
; FILING DATE: 03-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/863,703
; FILING DATE: 03-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 91 07566.3
; FILING DATE: 10-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Bieker-Brady, Kristina
; REGISTRATION NUMBER: 39,109
; REFERENCE/DOCKET NUMBER: 04585/00200Q
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 428-0200
; TELEFAX: (617) 428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 149:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1140
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-736-019-149

```

```

Query Match          9.6%; Score 95.4; DB 7; Length 1140;
Best Local Similarity 50.7%; Pred. No. 1.9e-17;
Matches 409; Conservative 0; Mismatches 361; Indels 37; Gaps 6;

```

```

Qy      194 GGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGGGGGCTGCAGCGCG 253
      ||| ||| | | | | | | | | | | | | | | | |
Db      11 GGGCGGCGAAAGCCGGGGGCTTGAAGAAGGACTCGCTGCTCACCCTGCGCCTGGGCGCCT 70

Qy      254 AGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAGCGCTACATCTTTT 313
      | | | | | | | | | | | | | | | | | | | |
Db      71 GGGGCCACCCCGCCTTCCCCTCCTGCGGGCGCCTCAAGGAGGACAGCAGGTACATCTTCT 130

Qy      314 TCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCCCCCTCGATACC- 372
      || ||||| || | | | | | | | | | | | | | |
Db      131 TCATGGAGCCCGAGGCCAACAGCAGCGGCGGGCCCGCCGCTTCCGAGCCTCCTTCCCC 190

Qy      373 -----AACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTGC 423
      | || | | | | | | | | | | | | | | | | |
Db      191 CCTCTCGAGACGGGCCGGAACCTCAAGAAGGAGGTGAGCCGGGTGCTGTGCAACGG-TGC 249

Qy      424 GCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAA 483

```

```

Db      250  |||  |  |||  ||||  |||||  |||  |  ||  |||  |||  |
          GCCTTGCCTCCCCGCTTGAAAGAGATGAAGAGTCAGGAGTCTGTGGCAGGTTCCAAACTA 309

Qy      484  TCGCTGAAGTGTGAGGCAGCAGCCGGAATCCCCAGCCTTCCTACCGTTGGTTCAAGGAT 543
          |||  |||  |||  |  |  |  |  |  |  |  |||||  |||

Db      310  GTGCTTCGGTGCGAGACCAGTTCTGAATACTCCTCTCTCAAGTTCAAGTGGTTCAAGAAT 369

Qy      544  GGCAAGGAGCTCAACCG---CAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAAG 600
          ||  |  ||  |  |||  |  |  |  |||||  ||  |  |

Db      370  GGGAGTGAATTAAGCCGAAAGAACAAACCAGAAAACATCAAGATACAGAAAAGGCCGGGG 429

Qy      601  AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660
          ||  |||  |||  |  |  |||  |  |  |||  ||  |||  |||  |  |||  |

Db      430  AAGTCAGAACTTCGCATTAGCAAAGCGTCACTGGCTGATTCTGGAGAATATATGTGCAAA 489

Qy      661  GCCGAGAACATCCTGGGGAAGGACA-----CCGTCCGGGGCCGGCTTTACGTCAACAGC 714
          |  |  |  ||  ||  ||  ||  ||  |  |  |  |  |  |  |  |

Db      490  GTGATCAGCAAACCTAGGAAATGACAGTGCCTCTGCCAACATCACCATTGTGGAGTCAAAC 549

Qy      715  GTGAGCACCACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAG 774
          |  |  |||  ||  |  ||  |  ||  ||||  |||  |  ||

Db      550  GCCACATCCACATCTACAGCTGGGACAAGCCATCTTGTCAAGTGTGCAGAGAAGGAGAAA 609

Qy      775  TCCTATTGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTC----- 830
          |  |  ||  ||  |||||  |||  |  |  |  |  |  |  |||

Db      610  ACTTTCTGTGTGAATGGAGGCGAGTGCTTCATGGTGAAAGACCTTTCAAATCCCTCAAGA 669

Qy      831  -----CTGCAAATGTCCAAATGGATTCTTCGGACAGAGATGTTTGGAGAAACTGCCTTTTG 885
          ||||  ||  |  |  |||||  |||  |||||  ||||  |||  ||

Db      670  TACTTGTGCAAGTGCCAACCTGGATTCACTGGAGCGAGATGTACTGAGAATGTGCCCATG 729

Qy      886  CGATTGTACATGCCAGATCCTAAGCAAAGTGTCTGTGGGATACACCGGGGACAGGTGTC 945
          |  |  |  |  ||  ||  ||||  ||  |  |  ||  ||  |  ||  |

Db      730  AAAGTCCAAA-----CCCAAGAAAAGTGCCCAAATGAGTTTACTGGTGATCGTGCC 781

Qy      946  AGCAGTTCGCAATGGTCAACTTCTCCA 972
          |  |  |  ||  |||  ||  ||||  ||

Db      782  AAAACTACGTAATGGCCAGCTTCTACA 808

```

RESULT 13

US-09-366-886-55

; Sequence 55, Application US/09366886

; Publication No. US20030040465A1

; GENERAL INFORMATION:

; APPLICANT: Gywnne, David I.

; APPLICANT: Mahanthappa, Nagesh K.

; APPLICANT: Marchionni, Mark A.

; APPLICANT: Bermingham-McDonogh, Olivia

; APPLICANT: Goldin, Stanley M.

; APPLICANT: McBurney, Robert N.

; TITLE OF INVENTION: USE OF NEUREGULINS AS MODULATORS OF

; TITLE OF INVENTION: CELLULAR COMMUNICATION

; FILE REFERENCE: 04585/041005

; CURRENT APPLICATION NUMBER: US/09/366,886

; CURRENT FILING DATE: 1999-08-04

; PRIOR APPLICATION NUMBER: US 08/341,018

Qy 775 TCCTATTGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTC----- 830
 | | | | | | | | | | | | | | | | | | | | | |
 Db 610 ACTTTCTGTGTGAATGGAGGCGAGTGCTTCATGGTGAAAGACCTTTCAAATCCCTCAAGA 669
 Qy 831 -----CTGCAAATGTCCAAATGGATTCTTCGGACAGAGATGTTTGGAGAAACTGCCTTTG 885
 | | | | | | | | | | | | | | | | | | | | | |
 Db 670 TACTTGTGCAAGTGCCAACCTGGATTCACTGGAGCGAGATGTACTGAGAATGTGCCCATG 729
 Qy 886 CGATTGTACATGCCAGATCCTAAGCAAAGTGTCTGTGGGATACACCGGGGACAGGTGTC 945
 | | | | | | | | | | | | | | | | | | | | | |
 Db 730 AAAGTCCAAA-----CCCAAGAAAAGTGCCCAAATGAGTTTACTGGTGATCGCTGCC 781
 Qy 946 AGCAGTTCGCAATGGTCAACTTCTCCA 972
 | | | | | | | | | | | | | | | |
 Db 782 AAAACTACGTAATGGCCAGCTTCTACA 808

RESULT 14

US-08-736-019-134

; Sequence 134, Application US/08736019

; Publication No. US20030207799A1

; GENERAL INFORMATION:

; APPLICANT: Goodearl, Andrew

; APPLICANT: Stroobant, Paul

; APPLICANT: Minghetti, Luisa

; APPLICANT: Waterfield, Michael

; APPLICANT: Marchionni, Mark

; APPLICANT: Chen, Mario

; APPLICANT: Hiles, Ian

; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR

; TITLE OF INVENTION: PREPARATION AND USE

; NUMBER OF SEQUENCES: 189

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Clark & Elbing LLP

; STREET: 176 Federal Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: U.S.A.

; ZIP: 02110

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

; COMPUTER: IBM Compatible Pentium

; OPERATING SYSTEM: Windows95

; SOFTWARE: FastSeq Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/736,019

; FILING DATE: 22-OCT-1996

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/471,833

; FILING DATE: 06-JUN-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/036,555

; FILING DATE: 24-MAR-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/965,173

Qy 544 GGCAAGGAGCTCAACCG---CAGCCGAGACATTGCGATCAAATATGGCAACGGCAGAAAG 600
 || | || | | || | | | || || | | |
 Db 377 GGGAGTGAATTAAGCCGAAAGAACAAACCAGAAAACATCAAGATACAGAAAAGGCCGGGG 436
 Qy 601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660
 || ||| |||| | | | || | | || | | || | || |
 Db 437 AAGTCAGGACTTCGCATTAGCAAAGCGTCACTGGCTGATTCTGGAGAATATATGTGCAA 496
 Qy 661 GCCGAGAACATCCTGGGGAAGGACA-----CCGTCCGGGGCCGGCTTTACGTCAACAGC 714
 | | | | | | | | | | | | | | | | | |
 Db 497 GTGATCAGCAAACCTAGGAAATGACAGTGCCTCTGCCAACATCACCATTGTGGAGTCAAAC 556
 Qy 715 GTGAGCACCAACCTGTCTCCTGGTCCGGGACGCCCCGGAAGTGAACGAGACAGCCAAG 774
 | | | |||| | | | | | | | | || || | ||
 Db 557 GCCACATCCACATCTACAGCTGGGACAAGCCATCTTGTCAAGTGTGCAGAGAAGGAGAAA 616
 Qy 775 TCCTATTGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCC--- 831
 | | || || || || || || || || || | | | | | |
 Db 617 ACTTTCTGTGTGAATGGAGGCGAGTGCTTCATGGTGAAAGACCTTTCAAATCCCTCAAGA 676
 Qy 832 -----TGCAAATGTCCAAATGGATTCTTCGGACAGAGATGTTTGGAGAACTGCCTTTG 885
 |||| | | | | |||| | | || || || || || ||
 Db 677 TACTTGTGCAAGTGCCAACCTGGATTCACTGGAGCGAGATGTACTGAGAATGTGCCCATG 736
 Qy 886 CGATTGTACATGCCAGATCCTAAGCAAAGTGTCTGTGGGATACACCGGGGACAGGTGTC 945
 | | | | | | | | || || || | | | | | | | |
 Db 737 AAAGTCCAAACCCAAG-----AAAGTGCCCAAATGAGTTTACTGGTGATCGCTGCC 787
 Qy 946 AGCAGTTCGCAATGGTCAACTTCTCCA 972
 | | | | |||| | | |||| | |
 Db 788 AAAACTACGTAATGGCCAGCTTCTACA 814

RESULT 15

US-09-366-886-3

; Sequence 3, Application US/09366886

; Publication No. US20030040465A1

; GENERAL INFORMATION:

; APPLICANT: Gywnne, David I.

; APPLICANT: Mahanthappa, Nagesh K.

; APPLICANT: Marchionni, Mark A.

; APPLICANT: Bermingham-McDonogh, Olivia

; APPLICANT: Goldin, Stanley M.

; APPLICANT: McBurney, Robert N.

; TITLE OF INVENTION: USE OF NEUREGULINS AS MODULATORS OF

; TITLE OF INVENTION: CELLULAR COMMUNICATION

; FILE REFERENCE: 04585/041005

; CURRENT APPLICATION NUMBER: US/09/366,886

; CURRENT FILING DATE: 1999-08-04

; PRIOR APPLICATION NUMBER: US 08/341,018

; PRIOR FILING DATE: 1994-11-17

; NUMBER OF SEQ ID NOS: 87

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 1193

; TYPE: DNA

; ORGANISM: Bos taurus

US-09-366-886-3

Query Match 9.4%; Score 93.4; DB 11; Length 1193;
Best Local Similarity 50.6%; Pred. No. 7.3e-17;
Matches 408; Conservative 0; Mismatches 361; Indels 38; Gaps 6;

Qy	194	GGGTGGCGTGTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGGGGGCTGCAGCGCG	253
Db	18	GGGCGGCGAAAGCCGGGGGCTTGAAGAAGGACTCGCTGCTCACCGTGCGCCTGGGCGCCT	77
Qy	254	AGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAGCGCTACATCTTTT	313
Db	78	GGGGCCACCCGCCTTCCCCTCCTGCGGGCGCCTCAAGGAGGACAGCAGGTACATCTTCT	137
Qy	314	TCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTGCCCCCTCGATACC-	372
Db	138	TCATGGAGCCCGAGGCCAACAGCAGCGGCGGGCCGCGCCTTCCGAGCCTCCTTCCCC	197
Qy	373	-----AACGGCAAAAATCTCAAGAAAGAGTGGGCAAGATCCTGTGCACTGACTGC	423
Db	198	CCTCTCGAGACGGGCCGGAACCTCAAGAAGGAGTCAAGCCGGGTGCTGTGCAACGG-TGC	256
Qy	424	GCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAA	483
Db	257	GCCTTGCCCTCCCCGCTTGAAAGAGATGAAGAGTCAGGAGTCTGTGGCAGGTTCCAAACTA	316
Qy	484	TCGCTGAAGTGTGAGGCAGCAGCCGGAATCCCCAGCCTTCCTACCGTTGGTTCAAGGAT	543
Db	317	GTGCTTCGGTGCAGACCAGTTCTGAATACTCCTCTCTCAAGTTCAAGTGGTTCAAGAAT	376
Qy	544	GGCAAGGAGCTCAACCG---CAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG	600
Db	377	GGGAGTGAATTAAGCCGAAAGAACAACCAGAAAACATCAAGATACAGAAAAGGCCGGGG	436
Qy	601	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Db	437	AAGTCAGGACTTCGCATTAGCAAAGCGTCACTGGCTGATTCTGGAGAATATATGTGCAA	496
Qy	661	GCCGAGAACATCCTGGGGAAGGACA-----CCGTCCGGGGCCGGCTTTACGTCAACAGC	714
Db	497	GTGATCAGCAAAC TAGGAAATGACAGTGCCTCTGCCAACATCACCATTGTGGAGTCAAAC	556
Qy	715	GTGAGCACCACCCTGTCTATCCTGGTCCGGGACGCCCCGGAAGTGCAACGAGACAGCCAAG	774
Db	557	GCCACATCCACATCTACAGCTGGGACAAGCCATCTTGTCAAGTGTGCAGAGAAGGAGAAA	616
Qy	775	TCCTATTGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCC---	831
Db	617	ACTTTCTGTGTGAATGGAGGCGAGTGCTTCATGGTGAAAGACCTTCAAATCCCTCAAGA	676
Qy	832	-----TGCAAATGTCCAAATGGATTCTTCGGACAGAGATGTTTGGAGAAACTGCCTTTG	885
Db	677	TACTTGTGCAAGTGCCAACCTGGATTCACTGGAGCGAGATGTACTGAGAATGTGCCCATG	736
Qy	886	CGATTGTACATGCCAGATCCTAAGCAAAGTGTCTGTGGGATACACCGGGGACAGGTGTC	945
Db	737	AAAGTCCAAACCCAAG-----AAAGTGCCCAAATGAGTTTACTGGTGATCGCTGCC	787

Qy 946 AGCAGTTCGCAATGGTCAACTTCTCCA 972
| | | | | | | | | | | | | |
Db 788 AAAACTACGTAATGGCCAGCTTCTACA 814

Search completed: January 14, 2004, 14:11:36
Job time : 392.249 secs